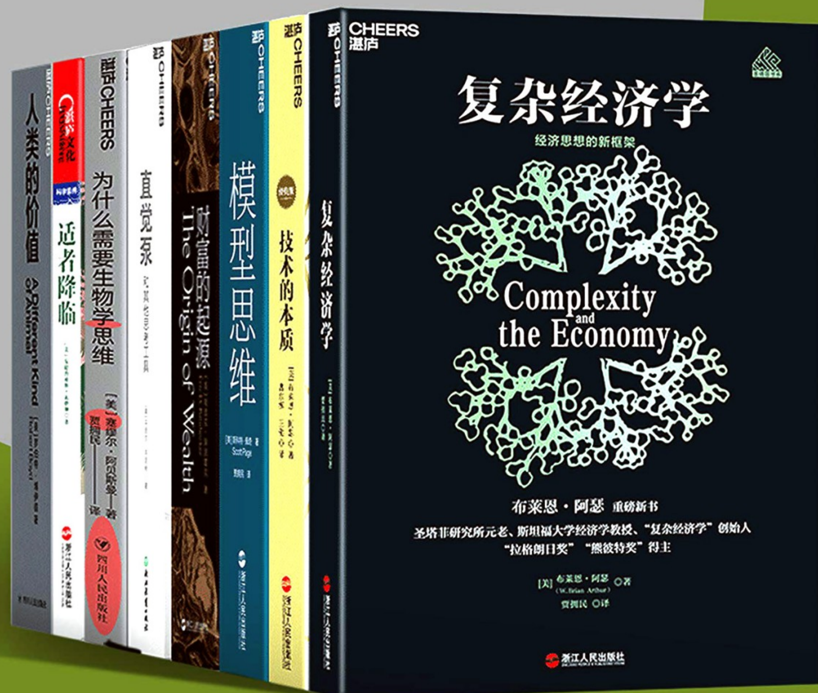
复杂性思维八部曲


人类的价值
适者降临
为什么需要生物学思维
直觉泵
财富的起源
The Origin of Wealth
模型思维
技术的本质
复杂经济学
CHEERS
湛庐
复杂经济学
经济思想的新框架
Complexity and the Economy
布莱恩·阿瑟 重磅新书
圣塔菲研究所元老、斯坦福大学经济学教授、“复杂经济学”创始人
“拉格朗日奖” “熊彼特奖”得主
[美] 布莱恩·阿瑟 著
贾拥民 译
浙江人民出版社

VISIT...

LANZAROTE Caliente.com

复杂性思维八部曲



人类的价值
A Different Kind of Animal
适者降临
为什么需要生物学思维
四川人民出版社
直觉泵
财富的起源
The Origin of Wealth
模型思维
技术的本质
复杂经济学
CHEERS
湛庐
复杂经济学
经济思想的新框架
Complexity and the Economy
布莱恩·阿瑟 重磅新书
圣塔菲研究所元老、斯坦福大学经济学教授、"复杂经济学"创始人
"拉格朗日奖" "熊彼特奖"得主
浙江人民出版社

版权信息

本套装纸版浙江人民出版社2019年12月出版

作者授权湛庐文化（Cheers　Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有•侵权必究

书名：复杂性思维八部曲

作者：安德烈亚斯•瓦格纳;布莱恩•阿瑟;塞缪尔•阿贝斯曼;埃里克•拜因霍克;斯科特•佩奇;罗伯特•博伊德;丹尼尔•丹尼特

电子书定价：770.99元

总目录

复杂经济学：经济思想的新框架
技术的本质：技术是什么，它是如何进化的（经典版）
模型思维
财富的起源
直觉泵和其他思考工具
为什么需要生物学思维
适者降临
人类的价值

与思想有关

business



复杂经济学

经济思想的新框架



Complexity and the Economy

布莱恩·阿瑟 重磅新书

圣塔菲研究所元老、斯坦福大学经济学教授、“复杂经济学”创始人

“拉格朗日奖”“熊彼特奖”得主

[美] 布莱恩·阿瑟（W. Brian Arthur）◎著

贾拥民 ◎译


浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE



复杂经济学

经济思想的新框架

[美] 布莱恩·阿瑟（W. Brian Arthur）◎著
贾拥民 ◎译

浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2018年5月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有•侵权必究

书名：复杂经济学：经济思想的新框架

著者：布莱恩·阿瑟

电子书定价：89.99元

W. Brian Arthur. Complexity and the Economy

Copyright © Oxford University Press 2015

古典经济学是人类农业文明的结晶，新古典经济学是人类工业文明的结晶，复杂经济学是人类信息文明的结晶！阿瑟的著作，为我们了解这个正在我们身边发生、并且将主导人类经济未来发展愿景的形态，提供了一种全面、直观和深刻的启迪。

叶航

浙江大学经济学教授，跨学科社会科学研究中心主任

越来越多的学者同意，传统经济学的思想框架和知识谱系，难以解释这个快速变化的世界。这个世界既不是市场失灵的问题，也不是政府失灵的问题，而是理论失灵。在阿瑟看来，经济学思想失灵的根源，在于过去还原论、确定性思想的禁锢。作为复杂经济学思想的创立者，阿瑟在这本书中用过去30年的历史文献，详细解读了复杂思想与经济学思想相互交融的心路历程，可谓字字珠玑，篇篇经典。

段永朝

苇草智酷创始合伙人，财讯传媒首席战略官

几百年后，很有可能我们会发现，以还原论思想为指导，试图建立一套类似牛顿力学的体系，去清晰刻画超级复杂的经济系统，是一场彻头彻尾的败局！希望幸运如你，在几百年前就读到《复杂经济学》这本书。

周涛

电子科技大学教授，成都市新经济发展研究院执行院长

自人类进入信息时代，世界经济发生了翻天覆地的变化，新现实呼唤着更具解释力的新理论。布莱恩·阿瑟的《复杂经济学》开拓了经济学研究的新领域，而其中的“复杂性思维”对于我们理解当今所处的时代环境具有深刻的启示。

吕琳媛

电子科技大学教授，阿里巴巴复杂科学研究中心副主任

布莱恩·阿瑟改变了我们看待经济现象的方式！经济系统的复杂性，一方面是由人们千差万别的预期所导致的，另一方面，收益递增规律也

决定了经济的未来进化。阿瑟用“复杂性思维”写出的这本《复杂经济学》，值得读者朋友们一读再读！

肯尼斯·阿罗
诺贝尔经济学奖获得者

布莱恩·阿瑟用他多年研究经济复杂性所写的文字，巧妙地搭建了一个经济思想的新框架，让他首创的“复杂经济学”一下子丰满起来。书中不时透出他那爱尔兰风格的幽默。阿瑟率先提出的“爱尔法鲁酒吧”问题，今天仍吸引着无数的后来者探究下去。如此有深度又好读的复杂经济学精品，真是独一无二！

约翰·霍兰德
密歇根大学计算机科学、工程学和心理学教授，“遗传算法之父”

布莱恩·阿瑟是推动经济思想改变的关键人物。《复杂经济学》把阿瑟对复杂经济的研究和他对技术本质的研究完美地结合在一起，既展示了他的创造性才华，也体现了他思维方式的组合进化。这本书为读者指明了经济学的未来发展方向。

大卫·柯南德尔
明德学院经济学教授

布莱恩·阿瑟是回报递增经济学的先锋人物。他是研究技术本质及其与经济关系的先锋。对于“复杂性”这一时髦而又难以理解的概念，《复杂经济学》给出了清晰的诠释。

马丁·舒彼克
耶鲁大学经济学教授，现代博弈论创始人

布莱恩·阿瑟关于技术本质的独到见解，会启迪所有的人，不论他们是技术的批评者、支持者，还是那些困惑不解的人。

凯文·凯利
《连线》创始主编

我们的Java，就是根据布莱恩·阿瑟的思想开发的。

埃里克·施密特

谷歌公司前董事长

理解“涌现秩序”

汪丁丁
北京大学国家发展研究院教授

就我所知，包括这本文集的作者阿瑟在内，研究“复杂现象”的学者们都承认，“复杂性”与“人”是最难定义的两大观念，如果坚持“内涵定义”而不是“外延定义”的话。例如，关于“人”，我只能跟随黑格尔说，人的本质是精神，而精神的本质是自由。关于“复杂性”（英文“complexity”有远比“complication”更复杂的涵义），我也只能说，它的本质（或“等价观念”）是“涌现秩序”，而“涌现”的本质是怀特海在《思维方式》里阐述的“过程”。我曾写过一篇晦涩短文《涌现秩序的表达困境》，几年前，阿瑟这本《复杂经济学》的译者贾拥民用来做了他另一本译作的序言。我把那篇文章也附在文后了。

阿瑟2015年3月4日在新加坡南洋理工大学“涌现模式研讨会”（Emerging Patterns Conference）上的演讲《复杂性与西方思想的迁移》（*Complexity and the Shift in Western Thought*），或许是他为自己这本《复杂经济学》提供的最新注释。此处，西方思想的核心是“科学”。阿瑟列出西方科学的四大基石（我更喜欢称之为西方科学的“心理基础”）：（1）秩序，或我更喜欢的表达，“秩序感”。（2）基于数学方程，或者，依照罗素和怀特海写作《数学原理》的初衷（将数学表达为逻辑的延伸），符合逻辑的表达。（3）可预测性，我认为这是最关键的。西方的崛起，借用赫拉利《人类简史》的概括，关键就是“知识与资本”的联姻。并且，在詹姆士《实用主义》哲学阐释之后，赫拉利正确地指出，培根“知识就是力量”的意思是，能增强人类力量的知识才是知识，而知识的真伪则是从属性的议题。从低等生命到高等生命，只要演化形成了在感觉神经元与运动神经元之间的“中介神经元”（inter-neurons）的网络，即可称为“中枢”的神经系统，行为就开始由“想象”（预测）与“行动”（实现）两个阶段组成。预测准确则行动效率高，也就是知识增强力量。（4）平衡状态或均衡状态，这是因为要预测就要观察（收集数据），而不平衡或非均衡的过程是很难观察的。

我同意张五常的判断，经济学，在社会科学诸学科中，最符合科学

的上列四性质。在同一演讲中，阿瑟指出，生物学是对上列四性质的最大挑战。因为，生物系统虽然有秩序感，却是开放的系统。并且，生物学一般而言不能表达为数学方程或模型。再者，生物演化通常不可预测。最后，生物过程不是均衡状态。由于一百多年来达尔文学说取得的辉煌成就，还由于基于牛顿力学的社会科学越来越难以适应互联网时代的复杂现象，西方思想正从牛顿的机械论模式向达尔文的演化论模式迁移。阿瑟这本《复杂经济学》，旨在澄清他多年甚至毕生努力要建立的“复杂现象的经济学”基本框架。

于是，阿瑟必须寻求一种新的表达。我读《复杂经济学》这本书，还没有见到这一令人期待的新的表达。直观而言，经济学的研究方法可概括为是“从本质到现象”的，我更喜欢说是“由内及外”的研究方法。经济学家根据观察得到一组内涵地定义他希望解释的经济活动的前提条件（偏好与约束），因为内涵定义是根据事物的本质而形成的，所以经济学家能从这组定义有逻辑地演绎出可在现实中获得验证（可证伪）的命题。与这一方法或许刚好相反的研究方法，是生物学的，可概括为是“从现象到本质”的，或者用我更喜欢的语言，就是“由外及内”的研究方法。与康德不能内涵地定义“人”的本质一样，生物学家不能内涵地定义“生命”的本质，于是生物学家只能从最表层的现象开始观察并确立自己的理解——即形成他所研究的生命过程的外延定义，并根据外延定义继续收集数据，以便得到更深入的理解和更深入的外延定义，从而更接近他所研究的生命过程的本质性的理解，逐渐观察并理解更深层的现象——这是一个不能穷尽的理解过程，以至于阿瑟指出，这一理解，它自身就是涌现秩序的一部分。读者必须时刻记着，这位阿瑟就是写了《技术的本质》的阿瑟。在学术思想史的视角下，阿瑟的思路更充分地表现于他关于技术之本质的论述中——我概括为两句话：（1）在演化视角下，技术就是有生命的；（2）在静态视角下，生命就是技术。

对于涌现秩序的表达困境，阿瑟必须解决的根本问题是：如何将经济学由内及外的研究方法与生物学由外及内的研究方法整合在同一理解框架内？

附言

涌现秩序的表达困境

涌现秩序（emerging orders）怎样表达，我相信，这一问题始终困扰着哈耶克，[1]也同样困扰当代研究复杂现象（包括演化社会理论）的学者，而且尤其因为意识到这一问题的深刻含义，与欧陆

和英美的其它思想传统相比，哈耶克更欣赏苏格兰启蒙时期的经验主义传统。[(2)]

首先，根据哈耶克的描述，将“复杂现象”区分于“简单现象”的认知特征在于，微观层次的行为主体不可能预见哪怕是服从最简单规则但数量极大的行为主体之间相互作用之后涌现出来的宏观秩序的样式，尽管他们可能事后理解这些样式。[(3)]其次，哈耶克相信，语言、人类的社会网络、每一个人脑内大量神经元的交互作用网络，这三类现象是复杂现象的经典案例。[(4)]最后，哈耶克试图论证：（1）存在“模糊型”与“清晰型”这两种不同的人类头脑或心智结构，（2）与清晰型头脑相比，模糊型头脑更富于原创性，因为，（3）脑内的创造性过程是一种复杂现象。[(5)]

从最简单的社会网络仿真不难看到，只要存在奈特所说的“不确定性”（而不是“风险”），并且只要相互作用的行为主体数目足够大，则不论行为规则多么简单，仿真研究者都不可能预见微观行为的宏观秩序。哈耶克试图论证的，在引入哪怕是最少量的奈特所说的“不确定性”之后，更可能获得清晰论证。为什么哈耶克必须假设有数量极大的行为主体服从哪怕最简单的行为规则？数量极大，意味着不可预期的小概率事件必定发生，相当于引入了奈特所说的“不确定性”。杨格（Peyton Young）在20世纪90年代得到的一个著名结果是，在随机过程的作用下，两个具有完备理性的博弈参与者相互观察对方策略的“学习过程”可以不收敛或收敛于错误的均衡。[(6)]因此，对于复杂现象之发生，与奈特所说的“不确定性”的作用相比，个体理性是否完备并不很重要，或者说，理性不完备只是奈特所说的“不确定性”的另一种表达，例如，表达为“C-D gap”的有限理性。[(7)]

在任何理解之前，先有表达。[(8)]怀特海的意思是，凡重要的，总要表达。面部表情、手势、声音、语言、文字、行动、情绪、群体行动和政治、战争和契约，凡有所表达的，都有重要性。绝大部分感受，不能获得而且可能永远不能表达。[(9)]许多被意识到的表达可称为“presentation”（呈现）或“representation”（再呈现）。虽然还有许多呈现或再呈现的重要性，从未被我们意识到。要概括无数具体的表达，荣格尤其重视“符号”，[(10)]他相信符号能够涵盖全部文化及其意义。关键是，人类或许永远只能理解符号的一部分涵义，[(11)]而由符号激发的精神过程的绝大部分是无意识的——个体无

意识和集体无意识。[12]脑科学进展到拉尔夫·阿道夫斯（Ralph Adolphs）提出“社会脑”概念以来，[13]符号互动论与社会交往理论获得了脑科学术语的表达。符号的涵义在社会交往中呈现于具体情境，孔子解仁，只在《论语》描述的那些具体情境之内阐释这一符号的涵义。完全脱离情境的符号，蜕变为“指号”（signs）。可以认为，符号是历史性的，而指号是逻辑性的。也因此，符号涵义是不可穷尽的。于是，符号涵义在历史过程中的呈现，与社会交往和人类的实践活动，有了密切联系，它们一起构成海勒女士阐释的“文化创造”过程。[14]

涌现秩序是历史过程，因为这些秩序只能通过“历史”（一系列的事件）呈现自身。斯密恪守“有限理性”原则，他从未将他对具体情境的政治经济分析扩展为一般原则，他始终对“体系之危险”（the danger of system）保持警惕。事实上，斯密的这一态度是苏格兰启蒙时期经验主义传统的心理特质。就斯密而言，这是一种源于斯多葛学派的身心修养。或许受到波兰尼的影响，[15]哈耶克对斯密的经验主义态度有一种远比同时代人更深切的理解。根据哈耶克的（或他尚未清晰表述的）理解，涌现秩序几乎是不能表达的，至少不能用统计方法来表达。（参阅《哈耶克文选》冯克利中译本“复杂现象论”这一章第4节的“统计学在处理模式复杂性上的不当”）。如果一颗清晰的头脑试图将某一新观念的全部内涵逻辑地表达到排除了任何隐秘知识的程度，那么，这一新观念的“新意”就将完全消失。因此，哈耶克相信，与创新过程相适应的是模糊型头脑。

了不起的阿瑟

张翼成
瑞士弗里堡大学终身教授，新经济体系的奠基之作《重塑》作者

很高兴看到布莱恩·阿瑟的大作《复杂经济学》由湛庐文化介绍给中国读者了。阿瑟是复杂科学圣地圣塔菲研究所的创始人之一，与其他批评主流经济学的经济学家不同的是，他曾在主流经济学领域获得过很高的地位与成就。比如他三十出头就被聘为斯坦福大学的终身教授了，而且他关于经济发展路径依赖的研究成果对当代的主流经济学有非常大的影响。最著名的诺奖获得者之一肯尼斯·阿罗（Kenneth Arrow）就曾经发表文章为阿瑟鸣不平：保罗·克鲁格曼（Paul Krugman）获奖的工作其实比阿瑟晚一年以上，有抄袭之嫌，但是阿瑟加入了反主流阵营，所以与诺奖失之交臂。好了，不扯远了，哪个圈子没有黑事呢？

我在22年前应邀去圣塔菲短住，最令我难忘的就是住在阿瑟最爱的爱尔法鲁酒吧屋后（一年后发生火灾，现在重建了），再一个就是结识了几位重量级学者，当时谈得最惬意的就是阿瑟。他当时介绍了他与那个酒吧同名的新作。该工作的新思路让我茅塞顿开，从此与经济现象结下了不解之缘。我回到瑞士一个月之后，正好有个新报名的博士生达米安·沙莱（Damien Challet）要开题。我就说咱们物理人应该把阿瑟的思想换个形式实现，然后两个月后我们就发表了《少数者博弈论》。没想到该文发表以后，数以百计的物理人也发现经济现象太诱人了，然后就有了经济物理这个新领域。换句话说，如果没有当年阿瑟的启蒙，经济物理等领域会像今天这样吗？

本书是阿瑟对他一生工作的总结，与其说是总结，倒不如说是反思。他清楚地认识到人们对经济社会现象的认知是很有限的，这一点他与库恩·索罗斯一致。他以此为出发点分析经济现象，尤其是企业如何创新发明。他与主流经济学的最大区别是信息极不完备，而且不可能有什么平衡态的。受他的思路启发，我们也在湛庐文化出版了新书《重塑》。与阿瑟不同的是我们仅仅关注消费市场，没有探讨更大范围的意义。但底层的哲理是一脉相承的。

相信阿瑟这本《复杂经济学》会惠及国内各行各业的读者大众，在

对主流经济学摧枯拉朽的批判中，在建立一个大的新理论框架过程中，此书提供了一个不可多得的指路明灯。

布莱恩·阿瑟访谈录

> 经济学领域正经历着100多年来最为深刻的变化。
>
> 埃里克·拜因霍克
>
> 经济学的新古典主义时代已经结束，取而代之的是复杂性时代。
>
> 理查德·霍尔特、小巴克利·罗塞尔和戴维·科兰德

问：经济学领域究竟发生了什么？

答：经济学正在发生重大变化。在过去的二三十年里，经济学家们认为，他们的标准方法，即新古典经济学，已经严重脱离了现实。新古典经济学假设人们是超理性的，在一个静态的、均衡的世界中进行决策。然而，经济学已经发生了分化，许多经济学家着手寻找更加现实的假设。因此，我们看到了行为经济学、收益递增经济学、进化博弈论的出现。复杂经济学也是如此。

问：那么，什么是复杂经济学呢？

答：复杂经济学是看待经济的一种完全不同的方式。复杂性其实是席卷了所有学科的一场运动，而不仅仅是一个研究课题。复杂系统是指组成系统的多个元素，要适应或响应这些元素自己创造的模式。复杂系统中的元素可以指元胞自动机中的元胞，或指交通系统中的汽车，前者会对相邻单元的状态做出反应，后者会对它前面或后面的汽车做出反应。当然，“元素”及它们所响应的“模式”在不同的情境下各不相同。但是无论如何，元素必须适应它们共同创造的世界，即总体模式。在这里，时间通过调整和变化自然而然地进入了系统：随着元素做出的反应，总量发生变化；而随着总量的变化，各种元素又重新做出反应。

经济中自然会出现复杂系统。经济行为主体，不管是银行、消费者、企业，还是投资者，都会不断调整自己的市场行动、购买决策、价格，并做出预测，以便适应所有这些市场行动（或决策、或价格、或预测）所共同创造的市场形势。所以，复杂经济学是看待经济的一种非常自然的方式，从某种意义上说，它已经存在200年了。复杂经济学实际

上是一种关于正在涌现的事物的经济学，它关注模式形成、结构变化、创新，以及永远的创造性毁灭的后果。

问：复杂性观点是否有逻辑基础？

答：当然有。正如我所指出的，经济中的参与者要不断地调整他们的市场决策、策略和预测，以适应这些行动（或决策、或策略、或预测）共同创造的市场形势。在这种情况下，理论经济学家很自然地就会去研究经济行为主体所创造的模式的展开。但是这显然是复杂的。因此，为了得到解析解，历史上的经济学选择了简化问题的道路。它反过来问了这样一个问题：什么样的行为导致了这样一种结果或模式，该结果或模式又导致了没有任何经济行为主体愿意去改变这种行为。换句话说，历史上的经济学提出的问题是，经济中的何种模式看上去是处于均衡状态的，即与创造它们的微观行为（或行动、或策略、或预期）是一致的。因此，一般均衡理论问的是：生产和消费的商品，在什么价格和什么数量上是与所在经济体的各个市场的价格和数量的整体模式相一致的，不会产生改变的激励？经典博弈论问的是：在给定对手可能选择的策略、行动或分配的前提下，什么样的策略、行动或分配是一致的，即对于一个行为主体来说，在何种标准下是最好的反应？这也算得上是研究经济学的一种比较自然的方法，却是“收益递减”的。

因此，很自然地，我们试图超越这种均衡方法，追问这样一个问题：经济行为主体的行为是如何不仅仅限于与它所创造的总体模式相一致的？或者，这个问题就是，经济行为主体的行动、策略或预期通常是如何对它们所创造的模式做出反应的，而且还可能内生地随它们创造的模式的变化而变化？换句话说，我们很自然地要问，当经济不处于稳态时，即不处于均衡状态时，经济是如何运行的。这就是复杂经济学。在这个更一般的层面上，我们也许可以推测，经济运行的模式可能会在足够长的时间后稳定下来，收敛到一个简单的、均质的均衡状态。但也可能是这样：它们表现出不断变化的、永远新奇的行为模式。它们可能会呈现出稳定状态下不会出现的新现象。

问：那么，是不是可以说复杂经济学和非均衡经济学密切相关呢？

答：是的。它们密切相关。事实上，我有时宁愿选择非均衡经济学这个术语。“复杂经济学”这个术语，是我在1999年为《科学》杂志撰写的一篇关于“复杂经济学”的文章中创造的。当时编辑要求我为这种新的经济学方法命名，所以我称之为“复杂经济学”。我稍微有点后悔。“非均衡”强调中断和破坏，这源于经济行为主体为了适应不断变化的情况而不

断进行的调整；“复杂”则强调经济行为主体对其他经济行为主体所导致的变化的反应。这两个概念密切相关。

问：复杂性与不确定性也是密切相关的，是不是？

答：是的。在复杂性方法中，你不能假定经济行为主体面临的所有问题都是确定的。这是因为，经济行为主体根本不知道其他经济行为主体可能会如何做出反应。他们不知道别人怎么看待同样的问题。因此，这是真正奈特（Knight）意义上的不确定性。这种不确定性意味着，经济行为主体需要“认知地”构建他们的问题，即必须“理解”问题，才能解决问题。所以这就将我们带入了认知经济学和行为经济学的世界。

问：你是怎么进入这个领域的？

答：在整个20世纪80年代，我一直在努力研究收益递增经济学。现在，收益递增经济学是复杂经济学的一个分支。当时我在斯坦福大学工作。1987年，肯尼斯·阿罗邀请我去圣塔菲研究所参加一个会议。当时，圣塔菲研究所才刚刚起步。一年后，我又回到圣塔菲研究所，领导一个名为“作为一个不断进化的复杂系统的经济”的研究项目。这其实就是圣塔菲研究所的第一个正式的研究项目。我们开始追问这样一个问题：经济失衡时会是什么样子？在这个研究项目中，我有一些非常优秀的同事：概率理论家戴维·莱恩（David Lane），物理学家理查德·帕尔默（Richard Palmer），理论生物学家斯图尔特·考夫曼（Stuart Kauffman），计算机科学家约翰·霍兰德（John Holland）。弗兰克·哈恩（Frank Hahn）、肯尼斯·阿罗（Kenneth Arrow）和汤姆·萨金特（Tom Sargent）则是访问研究员。从这个研究项目开始，复杂经济学成长了起来。当然，当时还有其他一些人也是这种新的经济学方法的建设者，在这里我想特别提一下以下学者的名字：彼得·艾伦（Peter Allen）、罗伯特·阿克斯特尔（Robert Axtell）、埃里克·拜因霍克（Eric Beinhocker）、乔希·爱泼斯坦（Josh Epstein）、多因·法默（Doyne Farmer）、阿兰·基尔曼（Alan Kirman）和李·特斯法齐（Leigh Tesfatsion）。现在，这种经济学研究进路已经蓬勃发展起来了，涌现了许多非常活跃的年轻学者。这一切都要追溯到圣塔菲研究所最早在这个领域做出的持续努力，而且很多方法也是在圣塔菲研究所最先出现的。

问：这种强调非均衡性和复杂性的观点，在经济学中是否已经有一段悠久的历史了？

答：在经济学中，这种思路其实已经有很长的历史了。我们正在探索的许多主题，如创新、中断、颠覆、在真正意义上的不确定性情况下的决策等，在熊彼特、凡勃伦、哈耶克、沙克尔，以及许多其他经济学家那里，都曾经被研究过。它们并不是经济学中的新问题。所不同的只不过是，我们现在可以更加严格地对这些主题进行研究。我们拥有了更多、更强大的工具，包括更复杂的概率理论，我们还可以在严格的控制下进行计算机实验。

问：你曾经谈到过，在经济学领域有两个非常重大的问题。它们是哪两个？

答：第一个问题是经济内部的配置（allocation）问题：在市场内部和不同市场之间，商品和服务的数量及它们的价格是如何决定的？对这个问题的回答，可以由一般均衡理论、国际贸易理论和博弈论等伟大理论来代表。第二个问题是经济内部的形成（formation）问题：经济最初是怎么出现的，又是怎么随着时间的推移成长起来并发生结构性的变化的？关于这个问题的回答，体现在关于创新、经济发展、结构变化，以及历史、制度和治理在经济中的作用等方面的思想上。配置问题现在已经被理解得比较充分，而且也已经高度数学化了。但是对于形成问题，经济学界现在的理解还很少，也几乎没有数学化。复杂经济学要研究的正是经济中结构的形成，因此它对形成问题和配置问题同样关注。

问：这是不是争议的焦点？

答：不，不再是了。复杂经济学是均衡经济学向非均衡情况下自然的延伸。而且由于非均衡包含了均衡，所以复杂经济学是经济学的扩展或一般化。这已经不再有争议了。复杂经济学的出现是不可避免的。这只是很多未来要完成的工作的开始。

问：如果复杂经济学真的像你说的那么重要，那么为什么我们在大学经济学系看不到复杂经济学呢？

答：不是的。其实我们已经看到不少了。但是，任何学科的改变，都需要一代人或更多时间才能完成。罗伯特·阿克斯特尔很喜欢举的一个例子是，博弈论花了大约四五十年的时间才算完全进入主流经济学殿堂。早在20世纪60年代就已经出现的行为经济学，也是到现在才刚刚开始登堂入室。从这个角度来看，复杂经济学仍有二三十年的时间才会进入主流。当然，对于我们这些研究者来说，也不是没有补偿，因为在一个全新的领域从事研究工作是最有趣的。我认为复杂经济学才刚刚起

步。

问：你刚才说复杂经济学是不可避免的，为什么这么说？

答：因为这不是一个喜不喜欢的问题。所有的学科都在经历重大转型：从将世界视为高度有序的、机械的、可预见的、在某种程度上静态的，转变为将世界视为不断进化的、有机的、不可预测的、处于永远发展中的。物理学、化学、数学、地质学等都是这样。经济学也不例外。虽然现在经济学相比其他学科稍微有点落在后面了，但是它终究是要追随时代精神的。

问：那么，复杂经济学是不是有什么“杀手级应用”，即没有它就不能做的事情？

答：有关复杂经济学的“杀手级应用”，我能想到的有两个。一个是我在20世纪80年代发展起来的收益递增经济学。它阐明了网络效应是如何导致“锁定”，或者说市场被一个或少数几个参与者统治的。这种工作是不能通过均衡经济学来完成的，因为这本身就不是一个均衡现象。现在，硅谷已经完全接受了这个理论，并根据它来运行了。

另一个“杀手级应用”是资产定价。复杂经济学从不假设存在一个（理性预期）均衡，也不打算去找到这样一个均衡。它假定，投资者并不知道市场是怎么运行的，必须自己去学会怎么做才是有效的，而这种学习本身就会改变市场。我们的实验结果重现了实际市场中出现的现象：技术交易（利用过去的价格模式来预测股票趋势，以获取利润）、价格和成交量高度相关、高波动性阶段和低波动性阶段交替出现（GARCH行为），等等。我们的理论解释了现实世界中的金融现象。



扫码下载“湛庐阅读”APP，
搜索“复杂经济学”，
获取中央电视台采访阿瑟的精彩视频。

目 录

推荐序一　理解“涌现秩序”
推荐序二　了不起的阿瑟
中文版序　布莱恩·阿瑟访谈录
引言　复杂性思维造就复杂经济学

在过去几十年里，完全理性、均衡、收益递减、独立的行为主体等关键假设，引起了很多经济学家的质疑。人们迫切需要全新的经济学观点，以应对非均衡、充满不确定性的经济世界。1999年，基于在圣塔菲研究所成长起来的“复杂性思维”，布莱恩·阿瑟首次提出“复杂经济学”这个概念。

结缘圣塔菲研究所
“人工股票市场”项目
“爱尔法鲁酒吧”问题
技术是如何进化的
复杂经济学的诞生
01　什么是复杂经济学　经济思想的新框架

复杂经济学以一种不同的方式来思考经济，它将经济视作不断进行自我“计算”、不断自我创建和自我更新的动态系统。它强调偶然性、不确定性、意义构建和“一切变化皆有可能”，是一门以预测、反应、创新和替代为基础的“动词”学科。复杂经济学正在取代新古典经济学而稳步走向经济学的中心。

经济与复杂性
内生的非均衡
建立非均衡状态下的理论
三种典型的非均衡现象
正反馈
经济的形成
泡沫和崩溃是市场的基本趋势
一门基于“动词”的科学
02　“爱尔法鲁酒吧”问题　归纳推理和有限理性

“爱尔法鲁酒吧”问题是个很有意思的决策问题。“很多人会去”的预期会导致几乎没人去，而“不会有人去”的预期则会导致许多人去。为什么理性预期会导致自我否定？现代心理学家认为，在复

杂的或不确定的情况下，演绎推理会显得“力不从心”，这时人们主要依靠归纳推理进行决策。

归纳思维
60人！归纳推理揭开的谜底
归纳推理的多彩世界
03　圣塔菲人工股票市场　内生预期的资产定价

资产定价理论认为：行为主体是完全异质的，他们的预期需要不断适应市场，而市场本身则是他们的预期共同创造的。实验表明，在归纳预期的股票市场上，市场行为显著地偏离了理性预期均衡，真实股票市场上的泡沫和崩盘等特征也都一一出现。

市场是理性的吗
为什么归纳推理是有效的
归纳预期的市场
两种市场体制的涌现
是普遍规律，还是人为假象
金融市场是个复杂系统
04　收益递增和路径依赖　技术竞争、正反馈及历史事件导致的锁定

复杂技术被采用后，往往表现出收益递增的特性。技术被采用得越多，获得的经验就越多，技术被改进得也就越多。当多种收益递增的技术为获得由采用者组成的市场而“相互竞争”时，某些看似微不足道的偶然事件可能变得相当重要，导致经济最终可能被“锁定”到某些不可预测的、较差的结果上。

简单模型：仅存在两种技术
一般框架：考虑随机性小事件
收益体制与路径锁定
动态收益递增会导致糟糕的“锁定”
05　经济中的过程与涌现　“复杂性视角”下的经济系统

复杂经济的显著特征告诉我们：传统经济学所使用的数学方法将面临极大挑战。把复杂性科学引入经济领域，将带来三个重大影响。所谓的“共同知识”，只能是通过具体的互动过程来获得经验，却无法用演绎推理方法获得。由行为主体之间重复互动模式所定义的网络结构，正成为新的研究热点。

复杂经济的6大特征
“复杂性视角”带来的三大影响
三个新热点

06　再好的经济和社会系统也会被“玩弄”　“压力测试”是防范操控行为的良方

再好的经济系统，也可能存在被他人利用的漏洞。我们可以借鉴工程上的失败模式分析法，对政策设计进行“压力测试”，模型中无须加入任何“剥削”元素。行为主体在做策略选择时，会发现某些行动特别有效，这时“剥削”行为就显露了身影。通过计算机的自动预警，我们就能采取相应的防范措施。

“剥削”的4种类型
“压力测试”的作用
用“涌现”发现新行为
自动预警
补上“失败模式反思”这一课
07　技术究竟是如何进化的　在实验室中观察到的组合进化

技术的进化机制，就是创造新的组合并选择那些有效的组合。所有技术都是从已经存在的技术中被创造出来的。复杂技术是涌现出来的！不过，这是通过先创造简单的、作为构件的技术来实现的。正如我们所预料的那样，只有少数技术被证明是创造“子孙后代”技术的关键构件。

电路设计的进化
新技术的涌现
技术的扩展
创造性毁灭的风潮
复杂技术是涌现出来的
08　技术进化所引发的经济进化　经济是从它的技术中涌现出来的

众多的技术集合在一起，创造了一种我们称之为“经济”的东西。经济从它的技术中浮现，不断地从它的技术中创造自己，并决定哪种新技术将会进入其中。每一个以新技术形式体现的解决方案，都会带来新的问题，这些问题又迫切需要进一步得到解决。经济是技术的一种表达，并随这些技术的进化而进化。

经济就是技术的一种表达
结构性变化
解决方案带来的新问题
09　复杂性的进化　是越来越复杂，还是随时可能坍塌

我们经常认为，随时间进化的系统，一般都会变得越来越复杂。实际上却并非如此。复杂性随系统进化而增加的机制有三种：

协同进化多样性的增加、结构深化和捕获软件。在这三种机制下，复杂性的增加都是间断性的和世代性的。由于前两种机制是可逆的，因此复杂性也可能随时坍塌。

机制1：协同进化多样性的增加
机制2：结构深化
机制3：捕获软件
复杂性的坍塌
10　认知科学　打开经济学黑箱的金钥匙

当人类行为主体面对复杂的或不确定性的决策问题时，他们在推理时所运用的不是演绎理性。那么，他们用的是什么呢？认知科学告诉我们，在这种情况下，我们会“联想地”进行思考：我们从经验中找到类似的情境，并用这些情境去拟合我们所面对的问题，然后从中得到一些启示。

心智是什么
心智是快速的模式完成器
认知过程建模
理论很重要，经验也很重要
认知真的有那么重要吗
11　确定性的终结　不确定性是经济世界的主旋律

很多年来，确定论和理性主义思想主导了理论经济学，这让许多经济学家深感不安。如果我们以完美的演绎推理预期采取行动，那么我们的预期就创造了正在试图预测的世界。但如果没有关于他人的预期的知识，任何一个行为主体的预期，从逻辑上看就不可能形成。实际上，经济中存在着根本的不确定性。

经济是确定的吗
“起飞时段选择”的困惑
“资产定价”的困惑
现实世界是这样的
结语　复杂的经济需要复杂经济学

在经济系统中，人类行为主体能够通过思考自己将要采取的行为可能带来怎样的结果，对自己的策略和预测进行调整，因此，完美的演绎推理就失效了。源于收益递增的正反馈是经济系统的重要特征。复杂经济学不是标准经济学理论的附属物，而是一个更一般的、超越了均衡观念的理论。

复杂经济学是一种超越均衡的经济理论

一只“轻推的手”
“归纳理性”的胜利
经济世界怎能少了“泡沫”和“崩溃”
经济是依赖于过程的、有机的和永远进化的
附录　未来的经济学原则
注释
译者后记
参考文献

COMPLEXITY

AND THE

ECONOMY

引言

复杂性思维造就复杂经济学

在过去几十年里，完全理性、均衡、收益递减、独立的行为主体等关键假设，引起了很多经济学家的质疑。人们迫切需要全新的经济学观点，以应对非均衡、充满不确定性的经济世界。1999 年，基于在圣塔菲研究所成长起来的“复杂性思维”，布莱恩·阿瑟首次提出“复杂经济学”这个概念。

任何一个学科都可能会进入这样一种“动荡期”：原先曾经被认定是理所当然的许多理论，似乎都不再那么可靠了；从事这个学科的人则公开探讨什么东西可以替代它们。经济学现在就处在“动荡期”中。一部分原因是发生于2008年的金融危机，但是对经济学传统理论的反思则在很早之前就已经开始了。过去30年，乃至更长的一个时期里，经济学家圈子内有一种氛围一直在非常缓慢地潜滋暗长：完全理性、均衡、收益递

减，以及永远只需面对有明确定义的问题的独立行为主体等关键假设，至少在某种程度上是不可信的、限制性太强的，或者说在一定意义上是迫不得已的。经济学家现在讨论得更多的是，行为理性、非均衡、收益递增，以及相互联系着的行为主体，他们决策时面对的是包含着根本的不确定性的问题。因此，除了标准的新古典主义的进路之外，经济学家还开辟了许多其他进路。

复杂经济学（**Complexity Economics**） 是一门超越了均衡层面的经济学理论，认为经济不是确定的、可预测的、机械的，而是依赖于过程的、有机的、永远在进化的。

其中一个进路就是“复杂经济学”（complexity economics）。我本人一直是这个领域的积极参与者。现在，我认为将自己以往撰写的文章归拢到一起，出版一本文集的时机已经成熟了。这本文集中的各篇文章都以经济与复杂性问题为核心，这些文章的写作时间跨度很大，早的可以追溯到20世纪80年代中期，晚的则止于当下此刻。这本文集的思想，与我以前出版的另一本讨论经济中的收益递增和路径依赖问题的文集，一脉相承。[1]

当然，所有这些“新”思想都不可能真的是“全新”的。在多年以前，甚至一个多世纪以前，许多经济学家就已经独立地以各种不同形式提出过了。但是对于这类洞见，以往一直缺乏处理它们的手段，如世界是不完美的，世界不是一架机器，整个世界无法被还原为一些只与实体的数量或“水平”的变化有关的简单方程式。不仅所需的技术还不成熟，而且所需的心态或心智结构也未具备。另外，也不存在一个以这些新思想为基础的、内在一致的经济学框架。

不过，在刚刚过去的几十年里，这种情况已经完全改变了。拼图的缺失部分已经逐渐填补上了，处理新的假设所需要的各种技术也慢慢发展成熟了。这些技术包括非线性动力学、非线性随机过程、基于主体的计算以及更一般的计算理论等。而且，人们的心态也变了。现在，在科学界，包括经济学界，研究者普遍认为，我们面对的世界不是一个完全有序的、原则上可以还原为几个数学方程式的系统；相反，在很大程度

上它是有机的和算法式的，即它是在原来就已经建立起来的系统基础上进化的，而且还是逐步进化的。由于这种种原因，经济学内慢慢地出现了一种基于上述更加现实的假设的研究进路。一个新的经济学框架正在形成。

复杂性（**Complexity**）是混沌性的局部与整体之间的非线性形式，由于局部与整体之间的非线性关系，使得我们不能通过局部来认识整体。

本书反映了我对这个新框架的发展做出的贡献。总的来说，这些文章的核心思想可以归结为：经济不一定处于均衡状态。在这个系统中，行为主体会不断地改变自己的行动和策略，作为对他们共同创造出来的结果做出的回应。这也就是说，在这个系统中，行为主体会不断地创造出一个“生态”来，而这个生态恰恰是他们自己必须与之相适应的。当然，这个观点的源头可以追溯到在20世纪70年代出现的所谓的“复杂性思维”（complexity thinking）。“复杂性思维”的成型，归功于布鲁塞尔、斯图加特和安娜堡的几个研究小组的工作，不过它最初的源头则来自斯坦福大学和麻省理工学院等大学的一些独立的研究者。不过，就经济学中的“复杂性思维”而言，却主要是在圣塔菲研究所（Santa Fe Institute）成长起来的。在20世纪80年代后期，圣塔菲研究所刚刚诞生，那里的一个研究小组就已经开始系统性地把经济视为一个不断发展的复杂系统来进行研究了。这个研究小组成立后的前两年，我本人领导了它的工作，而且自那之后一直与它保持着密切的联系。本书收录的文章表明，这些思想是如何发展起来并最终导致一门崭新经济学形成的。

本书中的各篇文章并不是一个“计划好的”研究过程的结果。它们是在一个很长的时期内断断续续写成的，而且受到了我的同事和圣塔菲研究所的一般思考方法的很大影响。其中有几篇文章曾经在著名的学术期刊发表过，其他一些文章则没有什么值得炫耀的“出身”。这些文章中有一部分是在圣塔菲研究所写的，其余文章则是在斯坦福大学写的。这些文章体现了我在写作时的思想，但是无法揭示这些思想为什么会产生以及它们是如何产生的。因此，为了帮助读者更好地理解它们，我先简要地叙述一下当初导致它们出现的背景和情境。

结缘圣塔菲研究所

我在本书中阐述的绝大多数思想都始于一个事件。

那是在1987年4月的一天，我正在斯坦福大学校园里走着，准备到自己的办公室去，突然，一个人骑着自行车围着我绕了个圈子，然后停在了我面前。那是肯尼斯·阿罗（Kenneth Arrow），他跨在自行车上，头上还戴着一个头盔。阿罗告诉我，9月份将会召开一个学术会议，让一群理论经济学家与一群自然科学家交流思想；与会的经济学家由他来确定，物理学家则由菲利普·安德森（Philip Anderson）来确定。会议将在圣塔菲研究所举行，那是位于洛基山脚下的一个刚刚启动的小型研究所。他问我想不想去参加，我立即回答："没问题！"尽管我还不知道自己在会议上要讲些什么，但是这个想法听上去就非常吸引人了。

几个月之后，会议如期在圣塔菲研究所举行。我发现，这是一个真正意义上的"重量级"会议，远远超出了我自己的想象。在与会的10位经济学家中，阿罗确定的人选包括了拉里·萨默斯（Larry Summers）、汤姆·萨金特（Tom Sargent）、何塞·沙因克曼（Jose Sheinkman）和威廉·布兹·布罗克（William Buz Brock）。而菲利普·安德森选择的10位科学家（物理学家）中，则包括了约翰·霍兰德（John Holland）[16]、戴维·吕埃尔（David Ruelle）、斯图尔特·考夫曼（Stuart Kauffman）和戴维·派因斯（David Pines）。这次会议是在圣塔菲研究所租来的一所修道院的小教堂中举行的，环境清雅静谧，会议的日程安排也非常宽松。每天上午，先由一位与会者发言，然后大家讨论；到下午，再由另一位与会者发言，然后大家一起讨论。所有这些学者，不仅要学习对方学科中解决问题的方法，还要搞明白对方学科是如何提出问题的，而且还有必要了解对方学科在考虑问题时的思维方式。许多问题，在经济学界通常不会有人提起，但是在这里都被物理学家们正儿八经地提出了：你们这群经济学家为什么要坚持完全理性？你们为什么要给出这么强的线性假设？当然，经济学家们也问物理学家：一个系统，例如自旋玻璃，明明还没有达到稳定状态，你们为什么就说问题已经"解决了"？共识是，在经济学和物理学中，都需要讨论混沌理论和非线性动力学，同时这两个学科都要对正反馈和相互作用进行建模。白天的会议结束后，与会者晚上还要继续讨论，不过不是集中到一起，而是两三人一组，互相交流想法和探讨问题。

这个会议是高强度的，既令人精神振奋，同时也令人疲惫不堪。当

10天的会议终于宣告结束时，与会者并没有得到任何明确的答案，但成果还是有的。在物理学家一方，他们理解了经济的极端复杂性：经济世界中的元素，即人，与物理世界中的元素，即“格子中的离子”不一样，因为人这种“元素”在决定下一步做什么的时候，不仅要依据自己和其他“元素”的当前状态，同时还要依据他们对这些“其他元素”，在给定他们自己可能做什么的条件下，可能会做什么样的推测。而在经济学家一方，也对现代物理学有了全新的真切感受：现代物理学强调相互作用和非线性；物理系统不但可以存在多个可能的最终状态，而且缺乏可预测性。这也就是说，经济学家也理解了物理学的复杂性。

复杂系统（**Complex System**） 是由大量组分组成的网络，不存在中央控制，通过简单运作规则产生出复杂的集体行为和复杂的信息处理，并通过学习和进化产生适应性。

会议结束后，我听到了一个消息：圣塔菲研究所接下来要有大动作。研究所的决策机构，即新成立的科学委员会决定，一旦会议收尾工作完成，就立即启动一个长期的研究计划，名为“经济可看作是进化的复杂系统”。他们邀请约翰·霍兰德和我明年来圣塔菲研究所主持这个研究项目。我在斯坦福大学刚好可以休一个年假，因此立即接受了；但是约翰却难以从密歇根大学脱身，因此只好婉拒。这样一来，就只能由我一个人来主持圣塔菲研究所的第一个研究计划了。该计划将从1988年8月正式开始执行。

回到斯坦福大学后，我的当务之急是，为这个新研究项目找到一批一流学者，组成一个高效的团队，并确定它的研究方向。我在这次会议中结识的一些学者可以参加这个研究计划。约翰·霍兰德答应来工作几个月，物理学家理查德·帕尔默也将会参与进来，而且他的时间还更充裕一些。斯图尔特·考夫曼则愿意长住在圣塔菲研究所。然后，我又从我自己的社交网络中寻找合适的人选。我邀请了戴维·莱恩和尤里·埃尔莫利耶夫（Yuri Ermoliev），他们两人都是出色的概率理论家。阿罗和安德森帮了我大忙。有几个人，我费尽唇舌也无法说服他们加入，而身为诺贝尔奖得主的阿罗和安德森却只要打一个电话，他们就马上兴冲冲地过来

了。至于研究方向，我一时之间还难以确定。一开始，物理学家默里·盖尔曼（Murray Gell-Mann）提出了一个建议：我们应该发布一份关于“构建一个全然不同的经济学”的宣言。但是，我对此还没有足够的信心；事实上，我这时还不能肯定我们究竟要设置哪些研究主题。虽然我已经在经济与复杂性这个方向上做了相当多的工作，但是现在可以供我们选择的主题范围更加广泛了。从这次会议来看，许多人倾向于把混沌理论作为一个核心主题，但是对我来说，这种想法没有什么吸引力。大体上，我认为我们更应该关注收益递增的问题，这是我更加熟悉的领域。相反，对于将物理学方法“移植”到经济学中来，和将非线性动力学应用于经济学研究的思路，我不是很感兴趣。但是，我们也许能够研究经济学中的计算这个有意思的课题。

1988年，当这个研究项目正式启动时，我们对研究方向的讨论更加深入了，不过仍然没有最后的定论。我们还在寻找前进的道路。我从圣塔菲研究所给肯尼斯·阿罗和菲利普·安德森打电话，征求他们的意见。他们又与这个研究项目的资助人、花旗银行的约翰·里德（John Reed）取得了联系。里德的回复是：想做什么都可以，只要研究的问题是关乎经济学的基础的，不是传统的、常规的就行。对于我及研究团队中的其他成员来说，这个“指示”不啻于一个美好的梦境。我们可以做我们想做的任何事！这极大地激发了我们的热情。在圣塔菲研究所，不会有来自本学科的同事在旁边盯着我们，更不会有人来问我们为什么做一些迥异于传统的事情。

事实上，在我们的小团队之外，圣塔菲研究所还有几位来自物理学或理论生物学的同事。斯图尔特·考夫曼就是其中一位，我们立即邀请他加入了我们的研究项目。圣塔菲研究所确实没有更多的渠道为我们提供研究人员。当时，它还只是一个寂寂无名的初创研究机构，偏居洛基山脉一隅，既没有学生，也没有教室；既没有院系，也没有学科。它自身也可以说是一个“实验品”，而且它的创办者还强调，就是不能有学科之分。

我们不停地讨论，主要的讨论场所就是那个修道院的厨房。我记得，有一天清晨，考夫曼说，你们经济学家为什么只研究均衡时如何如何？如果去研究非均衡经济学又会怎样？很可能，所有经济学家都像我一样，脑海中曾经闪过这个问题，却从来没有真正认真地去思考过。无论如何，当考夫曼提出这个问题时，我呆住了，其他经济学家也一样。我找不到合理的答案。它与下面这个问题属于同一类问题：如果不能考

虑重力，那么物理学将是什么样子？作为一个思想实验，这个问题当然是可以想象的，但它确实可以说是一个奇怪的问题。我们在考夫曼提出的这个问题上卡壳了，于是我们将它先放到一边，继续寻找前进的方向。

不过到了那个时候，我自己感兴趣的一个方向有一半已经成型了。这个方向是在前一年的会议上浮现出来的。就在会议的第一天午餐后的讨论中，约翰·霍兰德描述了他对"分类器系统"（classifier system）的研究。这里所说的"分类器系统"，大体上就是指这样一种系统，它由一系列描述"条件-行动"（condition-action）的规则连接而成。举例来说，一个规则可以是，如果系统的环境满足条件A，那么就执行动作R；另一个规则可以是，如果系统的环境满足条件D，那么就执行动作T；而第三个规则可以是，如果A为真，同时R动作被执行为假，那么就执行动作Z；诸如此类。行为者所采取的行动不但会改变环境，而且还会改变系统的整体状态。这样一来，只要将一系列这种"如果……那么……"规则串联在一起，你就可以让系统"识别"不同环境，并适当地执行特定的动作，就像大肠杆菌能够"识别出"环境中的葡萄糖梯度，并且朝适当的方向游动一样。此外，你还可以让系统从某些不那么好的规则开始，并用随着时间流逝而发现的更好的规则来替换那些不那么好的规则。这就是说，系统可以学习和进化。

当听霍兰德谈到这些时，我就已经非常兴奋了。我环顾四周，想看看同一房间内一起参加讨论的其他经济学家是不是也有我这种感受。没有任何证据表明他们有着跟我一样的感受。事实上，其中有一位经济学家甚至利用这个时间在午后小睡。不知何故，我有一种越来越强烈的感觉：在某种程度上，这可能就是我们最终的答案，而我们必须做的就是找到适当的问题。可以这么说，霍兰德正在描述的是"智能体"（intelligence）这样一个框架，或者说适当的行动是可以在系统内自动进化的。稍后，我向霍兰德详细请教了他的思想。1987年，我们在圣塔菲研究所的同一所房子里一起住了两个月，并讨论了好几次，但是我们两人都未能阐明如何将这些思想直接与经济学联系起来。

不久之后，我不得不返回斯坦福大学，因为我要在那里讲授发展经济学的课程。渐渐地，我形成了一个想法：我可以与约翰·霍兰德设计一个原始的人工经济，在我的计算机上运行。这个人工经济将使用他的学习系统，来生成越来越复杂、互为基础的行动规则，以此来模拟一个经济体从最简单的原始形式开始，发展为复杂的现代形式的进化过程。在

我想象出来的图景中，这个微型经济体及其迷你经济主体，将“安居”在我办公室角落的一台电脑里，自发地进化。只要我按下回车键，整个进化过程就开始了，然后我就可以去做其他事情了。也许几个小时后，我一回来就叫我的同事：“快来看看这些经济主体的行为，他们正在用黑曜石交换羊毛呢！”让计算机一直运行下去，一天后，也许我就会看到，为了进行交易，货币已经进化出来了，而且原始的银行也出现了。再过几天，股份公司也会出现。再到后来，我们将会看到中央银行、工会及工人偶尔举行的罢工、保险公司，甚至可以观察到期权交易的行为。这无疑是一个雄心勃勃的研究思路。我打电话将它告诉了霍兰德。他非常有兴趣，但是那时候，他和我都不知道到底怎样才能实现这个人工系统。

那时的情况就是如此。直到1988年6月的夏天，我们的研究方向仍然没有最终明确下来，但是研究项目已经启动了。不久之后，霍兰德和我再次在圣塔菲研究所聚首。我非常渴望对这种能够以某种形式自我进化的经济体进行研究。有一天，我和霍兰德在峡谷路一家名为“宝贝”（Babe）的餐馆吃午饭，他问我这个想法是怎么来的，如何才能将它落到实处。我告诉他，我发现，要让整个经济系统自发进化确实很难，但是还有一个更加简单的思路也许是可行的。我们可以模拟股票市场，而不直接模拟整个经济的进化。这个股票市场将是完全独立的。它将存在于计算机上，并且只有很少几个经济主体。更准确地说，“他们”是电脑化的投资者，每个投资者实际上是一个计算机程序。这些经济主体（即投资者）将会买入或卖出股票，试图发现涨跌趋势，甚至进行投机活动。我们可以从最简单的经济主体开始，然后允许他们通过运用霍兰德的“条件-行动”规则进行学习，这样他们就会逐渐变得聪明起来。我们可以对这个“股票市场”的行为进行分析，并将这些结果与真实市场的结果进行比较研究。霍兰德非常喜欢这个想法。

“人工股票市场”项目

入秋之后，一个基于计算机的股票市场模型的研究项目正式启动了。我们已经确定，我们这个“市场”中的“投资者”都是一些计算机程序，因此“他们”可以在一台放在我的办公桌上的电脑里，做出各种投资决策并进化。这个思路是很清晰的。但是我们经过多方尝试，依然未能成功地将股票市场行为归结为一系列的“条件-行动”规则。这样一来，我们的模型就显得过于特别（ad-hoc）了，然而我认为它还不够简洁、不够清晰。这时候，汤姆·萨金特正好从斯坦福大学来圣塔菲研究所访问，

他建议我们直接以罗伯特·卢卡斯（Robert Lucas）于1978年提出的股票市场模型为基础来构建我们的模型。这个建议是可行的，而且那样的话，我们的模型也会很简洁，同时也容易实现。当然，卢卡斯的模型是一个数理模型，它是用方程式表示的。为了便于分析，他的模型中的所有投资者都是同质的：他们以相同的方式对市场信号做出反应，而且正确程度都是同样的平均水平。卢卡斯已经以数理方式证明，股票价格将随着最近的收益序列而波动。

> **涌现（Emergence）** 就是指系统中的个体遵循简单的规则，通过局部的相互作用构成一个整体的时候，一些新的属性或规律就会突然一下子在系统的层面诞生。

相比之下，我们的投资者对股票市场的看法会有所不同，而且他们必须学会去判断：在股票市场上哪些东西是有效的、哪些东西是无效的。我们可以利用约翰·霍兰德的方法来实现这一点。我们这个人工股票市场中的投资者，将制定自己的“市况-预测”（condition-forecast）规则。例如，如果价格在过去三个交易期内都在上升，并且成交量下降超过了10%，那么预测下一个交易期的价格将会上升1.35%。我们还允许，每个投资者可以运用好几个这种规则，即复合规则或多重假说，它们都可能是适用的。关键是，在任何时候，他们都将按照最近被证明是最准确的那个规则来采取行动。当然，每个投资者运用的规则或假说是因投资者而异的。每个投资者以随机选择的规则开始，如果规则无效就会被抛弃，如果成功则可以被重新组合，从而潜在的新规则将会涌现出来。从一开始时，我们的投资者可能并不是非常“聪明”，但是随着时间的推移，他们会发现哪些规则是有效的，而且变得更加聪明起来。当然，这样也就改变了市场，于是我们的投资者可能不得不随时调整规则并发现新规则。

股票市场模型的第一个版本是在一台麦金塔计算机上运行的。负责编程的是物理学家理查德·帕尔默，他所采用的编程工具是Basic语言。我们的第一个目标是，努力让人工模拟股票市场运行起来，即让我们的投资者（计算机程序）根据他们自己当前对市场的了解，在市场上进行

出价和要价，并让整个市场正确地出清。这些目的很快就实现了。令人沮丧的是，经过初步观察，我们并没有发现这个人工股票市场的结果与标准的经济学模型的结果有任何不同之处。但是，随后通过更加细心的观察，我们注意到了真实的市场现象也出现在了这个人工股票市场中：小小的泡沫、小小的崩溃，以及价格和成交量的相关性，还有高度波动的交易时段与相对静止的交易时段的交替出现。我们的人工股票市场呈现出了现实世界中的现象，而标准经济学模型，即经济主体是同质、拥有理性预期能力的模型，是不可能得到这种结果的。

我们的人工股票市场，可以再现标准经济学模型不能再现的真实现象，这令我非常兴奋。在那个时候，我们就已经知道，我们做的是全然不同于传统的事情。我们正在模拟的是这样一个市场：个人采取行动相互竞争，并从他们的行动中进化出了一个“生态”。这个生态是个体之间的交互行为创造出来的，这是不可能通过基于方程式的标准经济学方法得到的。如果预测规则是由特定的条件触发的，并且每个投资者所采取的预测规则都是不同的，那就太复杂了，是无法用标准经济学方法来研究的。而且，我们这个人工系统与1986年前开始出现的其他基于规则的计算机仿真模型也不一样。在那些模型中，一方面，规则的数量很少；另一方面，所有规则都是事先固定好的。计算机仿真实验的目的只是让各种规则相互竞争，从而完成对它们的测试。我们的规则则是可以改变的，也会变异，甚至会“变得更加聪明”。我们有了一个真切的感觉，这个计算机仿真模型将使我们能够避免标准经济学模型，或者通常基于规则的系统过分简化的毛病。不过，我们并不认为我们的模型是对整个市场的模拟。我们只是把它当成一个实验室的实验来看的。利用它，我们可以先构建出一种基本情境来，然后系统地进行调整，来探索各种可能出现的结果。

对于这类研究，我们还没有一个恰当的名字。有一个阶段，我们称之为“基于元素的建模方法”，以便与基于方程的建模方法区分开来。大约三年后的1991年，约翰·霍兰德和约翰·米勒（John Miller）合作发表了一篇文章，系统地讨论了基于“人工适应主体”（artificial adaptive agents）建模的方法。[2]而在经济学界，人们则采用了“基于主体的建模方法”这个名字。

在圣塔菲研究所，经济学研究项目启动后的第一年里，我们还考虑了其他的一些问题。我们的思路是，不要去试图提出一个全新的经济学一般方法，尽管萨缪尔森（Samuelson）和其他的经济学家在几十年前

就是这么做的。相反，我们将重新考察一些已知的经济学问题，一些经济学中已经有了定论的问题。我们要从我们自己的、不同于传统的角度去重新解决这些问题。于是，约翰·鲁斯特（John　Rust）和理查德·帕尔默开始以这种方式研究双向拍卖市场。戴维·莱恩和我则用随机模型研究信息扩散（information　contagion）问题，而信息扩散模型是社会学习模型的早期版本。我原本以为，收益递增和正反馈的思想可以定义经济学研究项目启动后第一年的工作。但是他们并不这么认为。从根本上看，真正能定义这一年工作的是约翰·霍兰德提出的适应和学习的思想。我还以为，我们的进展可能是比较缓慢的，因为我们似乎没有得到太多的结果，但是到了第一年结束时，即1989年8月，肯尼斯·阿罗却告诉我们，与20世纪50年代考尔斯基金会资助的研究项目在最初的几年相比，我们的项目推进得更快，也更被人们所认可。

1990年，我离开圣塔菲研究所，回到了斯坦福大学。圣塔菲研究所的研究项目则由其他人接手。在整个20世纪90年代和21世纪初，这个项目在不同的项目主任的领导下，一直在继续着，而且相当成功。由于各位主事者有着各自的兴趣，这个项目下的研究主题有几年显得比较反传统，有几年则显得相对正统。1995年，我又回到了圣塔菲研究所，继续参与这个研究计划，历时5年之久。

收录在本书中的大多数经济学文章，都源于圣塔菲研究所经济学研究计划的头10年。关于我们的人工股票市场，我们先于1992年在《理学A》（*physica　A*）杂志上发表了一篇文章，后来又于1997年发表了另一篇文章。本书收录的是后一篇文章。这篇文章引起了广泛的关注，而且对“基于主体”的经济学研究产生了持久的影响。

“爱尔法鲁酒吧”问题

发表于1994年的另一篇文章也受到了高度关注（收录于本书第2章）。我把这篇文章称为“爱尔法鲁酒吧”问题，因为它的灵感来自于我对圣塔菲研究所旁边的一家名为“爱尔法鲁”（El　Farol）的酒吧的观察。每个星期四晚上，爱尔法鲁酒吧都有爱尔兰音乐专场，往往会爆满。如果酒吧里的人不太多，那么待在那里就很令人愉快；但是如果酒吧过于拥挤，那么它能够给你带来的乐趣就会少很多。我猜想，在某一个特定的晚上，如果每个人都预测有许多人会来，那么他们就不会来，这样的结果就会否定预测；如果每个人都预测有很少人会来，那么他们就会

来，这样的结果同样会否定预测。这就是说，理性预测（即理性预期）在这种情况下是自我否定的，因此能够正常发挥作用的理性预期就无法形成。

我很好奇的是，人工系统中的行为主体在面对这种情况时的行为会是怎样的。于是在1993年，我编写了一个程序，然后写了一篇文章。这篇文章发表在《美国经济评论会议文章》（*American Economic Review Papers and Proceedings*）上。读到这篇文章的经济学家几乎不知道该说些什么、做些什么。但是，它引起了著名物理学家、自组织临界理论的提出者帕·巴克（Per Bak）的注意。帕·巴克把它传真给了很多同事，于是突然之间，“爱尔法鲁”成了物理学家圈子中一个广为人知的名字。三年后，弗莱堡大学的物理学家达米安·夏利和张翼成，在我这个“爱尔法鲁”问题的基础上提出了少数者博弈（Minority Game）模型。[3]现在，“爱尔法鲁”问题和少数者博弈已经得到广泛而深入的研究，见诸于学术期刊的文章已经达到了几百篇之多。

技术是如何进化的

到了1997年，我又有了一个新的想法，不过这个想法与圣塔菲研究所的经济学研究计划没有直接的关系。我对技术产生了浓厚的兴趣。一开始，这种兴趣本身令我很困惑。尽管我的早期教育背景是工程技术学，但是这种兴趣的出现仍然显得有点奇怪，因为对技术魅力的着迷似乎与我的主要兴趣，即对经济学或复杂性的兴趣无关。我以前也研究过与“技术”相关的问题，但是当我开始深入探究各种技术为了被采用而展开竞争这种思想时，这种兴趣却消失了。我注意到，我所观察的所有技术，没有任何一个是纯粹地产生于某种灵感的。所有新技术都是原先已经存在的技术的组合。举例来说，激光打印机是由计算机处理器、激光器和静电复印技术组合而成的，计算机处理器引导激光器在复印机的硒鼓上“涂上”字母或图像，然后再复印出来。

我还意识到了其他的一些东西。1992年，出于好奇心，我一直在研究喷气式发动机。我希望搞清楚的是，在一开始出现的时候，喷气式发动机是那么简单，而在接下来的短短二三十年内，它们却变得如此复杂了，这是为什么。当时，我还一直在学习C语言编程。在我看来，用C语言编写的程序的结构与喷气式发动机的结构基本上是相同的。它们都有一个中央功能模块，还有一些支持这个核心模块的子模块，后者是用来

正确地设置和管理核心模块的。就某项给定的技术而言，随着时间的推移，如果添加上一些突破了以往的物理界限或能够更好地解决问题的子技术，那么就可以进一步“挤压”核心模块，让它发挥更大的性能。因此技术在刚出现时是简单的，但是随着技术的进化，更多的零件和子模块将不断地添加进来。1993年，我在《科学美国人》杂志上发表了一篇文章，讨论了各种系统为什么会越来越倾向于复杂和精致的问题。[4]

逐渐地，我心中有了这样的一个感觉：对于技术，我有一些一般性的东西可说。是的，我应该可以提出一个关于技术的一般理论。在此之前，我已经广泛阅读了与技术问题有关的文献，并且下定决心要学习和掌握多种技术。事实上，我对20来种技术都有很好的了解。到最后，我研究过的技术不仅包括了前面提到过的喷气式发动机，而且也包括了早期的收音机、雷达、蒸汽机、信息包交换技术、晶体管、麦克风、计算技术，甚至还包括了其他一些奇怪的“技术”，如青霉素等。这项研究工作的大部分是在圣塔菲的圣约翰学院（St. John's College）图书馆中完成的，也有一部分则是在我现在工作的施乐帕洛阿尔托研究中心（Xerox Parc）完成的。我总结出了技术形成和出现的一般模式。所有技术都是对现象的刻画和运用。归根结底，技术无非是用于实现人类目的的现象。现象是可以划分为不同“家族”的，如化学方面的、电子方面的、基因组方面的，因此也就形成了化学技术、电子技术、生物技术等这样的技术群。

而且从总体上看，如下这一点是显而易见的，不仅像喷气式发动机这样的单个技术会在其生命周期中不断进化，所有单个技术的总体集合在如下意义上也是进化的：在任何一个时刻，所有技术就像所有物种一样，都可以通过一条“祖先线”追溯到某项早期技术。但是，技术进化的基本机制却不是达尔文式（Darwinian）的。新技术并不是随着早期技术的微小变化的不断积累而出现的。很显然，喷气式发动机肯定不是从空气活塞发动机的微波变化的积累中产生的。新技术来自于对早期技术的组合或集成，人类的想象力和聪明才智也在其中发挥了重大作用。这就是说，有一个不同于达尔文式的技术进化机制。我称之为通过组合而进化（Evolution by Combination），或组合进化（Combinatorial Evolution）。

当然，这种进化机制同样也存在于生物进化当中。生物进化中出现的最重大的转变，绝大部分都是组合式的。单细胞生物以组合方式成为多细胞生物，原核生物以组合方式成为真核生物。但是，这种事件的发

生频率是极低的，要等上很多亿年才会出现一次。生物的日常进化机制是达尔文式的微小变化的积累，以及对这种变化的差异性选择。相比之下，在技术的进化中，主要的进化机制是组合式的，当然在新技术出现之后，也不排除达尔文式的技术进化，即微小变化的积累。

经过一个时期的研究，我相信对于技术是如何出现的，作为整体的技术集合是如何进化的，我已经洞若观火了。接下来，我想试一试，能不能在实验室里或计算机上实现这种技术的进化。2005年前后，我在富士施乐的FXPAL实验室工作，它实际上是富士施乐公司在帕洛阿尔托建立的一个智库。在那里，我遇到了计算机科学家沃尔夫冈·波拉克（Wolfgang Polak）。我提议我们可以进行这样一个计算机实验：把一些原始技术放到一起，组成一锅“技术之汤”，在这锅“技术之汤”中，各种技术可以随机地进行组合。由此得到的技术组合，即某种潜在的新技术，如果没有用就抛弃；如果有用就保留下来，并加入到那锅“技术之汤”用于进一步的技术组合。这样一个能够以上述方式不断创建技术集成的系统，能不能以“自展向上”（bootstraps）的方式使技术从简单进化为复杂？我们尝试了好几个系统，但是都没有成功。就在那个时候，我们读到了一篇非常漂亮的文章，它是由理查德·伦斯基等人撰写的，并发表在《自然》杂志上。[5]伦斯基和他的同事利用遗传算法进化出了数字电路。这篇文章启发我们，数字技术似乎是一种天然培养基。如果你把两个数字电路组合起来，你就可以得到另一个数字电路；当然，新的电路可能是有用的，也可能是没有用的。

然而，我们的实验并不容易。努力尝试了几个月之后，波拉克终于让我们的系统运行起来了。它利用简单的电路“创建”出新的电路。我们先准备好一锅由最简单的二位与非（nand）电路组成的“汤”，这种电路是数字电路的基本构件。然后按下回车键，再等上20个小时，看看从这锅汤中，能不能创建出什么逻辑电路。结果很理想，我们观察到了各种各样的电路。首先形成的是一些最基本的电路。然后出现的是一些中等复杂程度的电路，如4位的等于（equal）电路、三位的小于（less than）电路，等等。到最后，还形成了8位异或（exclusive-or）电路、8位与（and）电路和8位加法器（adder）。如果不细想，你可能觉得这没有什么了不起的。但是不要忘记，一个工作正常的8位加法器其实非常不简单，即将8位的x加到8位的y上，产生9位结果的z，它拥有16个输入针和9个输出针。而拥有16个输入针和9个输出针的电路组合，其种数高达$10^{177\ 554}$之多，但只有一种8位加法器是能够正常工作的。这不是随机的。因为我们这个实验有25万个步骤，通过这么多步随机地发现一个

8位加法器的概率，是完全可以忽略不计的。因此，我们这个连续集成过程，能够把原始基本构件组合起来，生成有用的简单构件，然后再把这些简单构件组合起来，生成进一步的构件。这无疑是非常强大的。现实世界中的真实技术也是以这种方式进化的。技术已经以自展向上的方式，从少数技术发展到许多技术，从原始的简单技术发展为高度复杂的技术。

我们在《复杂性》杂志（*Complexity*）上报告了我们的实验。但奇怪的是，我们这篇文章基本没有受到什么关注，也没有多少人来评论它。原因何在？我的猜测是，因为它是一篇夹在裂缝当中的文章。它既不是关于生物进化的，也不是讨论遗传算法的；它不是纯技术的，也不是纯经济学的。而且，我们的实验也没有解决某个特定的问题。它只是生成了一个有用的电路工具箱或电路库，就像某种编程语言的开发者提供的有用的函数库一样。但是关键在于，这个工具箱或库完全是进化出来的，我认为这可以说是一个奇迹。我拥有电气工程学位，波拉克拥有计算机科学学位，如果你要求我们亲自动手去设计一个8位加法器，那么我们就不得不从数字电子技术从头开始设计。但是现在，我们设计出了一种算法，它可以通过进化自动设计出这种电路。我相信，这是一种非常卓越的思想。而且我还认为，在本书所收录的这些文章中，这一篇文章有特别重要的意义，它阐明了一种处在不断进行中的进化，还阐明了一个不同于生物进化的进化机制，即通过组合或连续集成的进化。

然而，不知怎么的，我内心深处总有一种想法挥之不去。我认为，所有这一切都必须与经济进化相适应。如果用更准确的表达，应该是经济从一开始是如何在这一切中形成的。在我研究技术的过程中，我逐渐意识到，一方面，经济创造了技术；另一方面，同时也是更加重要的一方面是，技术即我们用来满足人类需求的技术集合创造了经济。因此，经济不仅是技术的容器，而且还是技术的表达。随着技术的进化，以及全新技术的引入，经济必然会发生变化。经济做什么，要发生变化；经济如何做，也要发生变化。经济的整个安排上，包括它的制度和机构上都必须发生变化，以适应于新的“做事方式”。简而言之，经济结构必须发生变化。

我在2009年出版的《技术的本质：技术是什么，它是如何进化的》（*The Nature of Technology: What It Is and How It Evolves*）[17]一书中，报告了上面这些发现。这本书受到了专业工程师的热烈欢迎，并且已经被翻译成了多种语言。在本书收录的这些文章中，其中有几篇是在写作

那本书时的“中间产物”，有一篇则是那本书的一个组成部分。从开始到结束，我对技术的本质和进化的研究花了整整12年的时间。我发现这个课题非常令人着迷。特别感到惊喜的是，我发现了技术集合进化的机制，还认识到了技术是一种具有自身内在逻辑结构的事物。通过这些研究，我越来越相信，技术本身的每一方面都是复杂的、有结构的，就像经济或法律体系一样。当然，技术也是一个非常美好的事物。

复杂经济学的诞生

我的思想之旅是由多条研究路线构成的。这些研究路线，在当年的我看来似乎是截然不同的。但是到了后来，尤其是今天，当我回过头去重新思考它们，再结合我在圣塔菲研究所和其他机构的同事的工作来看，我却发现，从这些研究路线中，一种全新的经济学观点已经慢慢生成了，而且逐渐地浮出了水面。1999年，我在《科学》杂志上发表了一篇文章，对我早期在这个方面的思考进行了初步总结。[6]《科学》杂志的编辑坚持要求我给这种不同于传统的经济学观点取个名字。于是我称之为“复杂经济学”。现在回想起来，复杂经济学的特征是非常清楚的。经济不一定处于均衡状态，经济通常都是处于非均衡状态。经济行为主体不是全知全能和完全理性的，他们必须理解他们所处的情境，并且在这样做的时候必须搜寻适当的策略。经济不是给定的，不可能是一个简单的技术容器；技术形塑了经济，在这过程中经济的结构是会变化的。因此，经济是有机的、分层的，后一层形成于前一层之上；经济永远都在变化，永远都在呈现新异性；在经济内部，结构会浮现，在持续了一定时间后又会消融。我要强调的是，所有这一切绝对不只是一种诗意化、人文化的描述，而是一种严谨的经济学观点，它能够被严格定义、被精确地探究和分析。

我经常被人问起，这种全新的经济学如何适应于标准的经济学分析？它难道不是标准经济学的一种简单的变体吗？用经济学家理查德·布朗克（Richard Bronk）的话来说，它会不会被“无缝”吸收进新古典经济学框架中？对于这些问题，我的答案是否定的。这种不同的经济学框架，既不是单纯利用计算机来进行基于主体的建模，也不是将对技术变迁的更加深刻的理解，加入内生经济增长模型中去。这种经济学要做的事情、所关注的东西、所依据的基本假设，都不同于标准经济学，它特别关注非均衡。这就是说，除了采用的方法不同之外，要解决的问题是不同的，解的概念本身也是不同的。

要理解这一点，要先明确一个有效的途径，标准的新古典经济学源于一种特殊的观察和认识世界的方式。新古典经济学继承了启蒙运动的思想，即我们观察到的混乱无序的世界只是表面现象，背后隐藏着秩序、理性与完美。它还继承了19世纪末的物理学观念，尤其是这种观念：大量相互作用的同质元素，可以通过简单的联立数学方程式一次性地全部分析清楚。到了20世纪中叶，这种观念在经济学中导致了如下这种愿景：经济学理论的核心，可以简单地用数理方程表示的定律来刻画，从而实现公理化。经济学理论的其他一些部分，如宏观经济学或制度经济学理论，也许不得不暂时先放一放，但是经济学理论的核心则肯定是可以被"规训的"，即被有序化和规律化的，并还原为数学。

这个研究纲领充其量只获得了部分成功，当然这部分成功的意义也不容小觑。其作用体现在，一方面，经济学的"门户"得到了清理，以前已经被接受为"经济学理论"的大量松散的、草率的论断被排除掉了；另一方面，人们对市场和资本主义制度的内在优势更加尊重，理解也更加透彻了。但是，我相信这种努力也导致了思想的僵化，还导致了一种貌似正义、实为党同伐异的判断标准。某些东西可以被承认为经济学理论，而另一些东西则不被允许，最终的结果是经济学成了一个无法接纳其他思想的封闭体系。由此进一步导致了政治、权力、阶级、社会、根本的不确定性、创造生成和发展对经济的影响，全都被"关在了经济学殿堂的门外"。最终结果则事与愿违，这个研究纲领，至少它的超理性版本，已经失败了。如果进行波普尔式（Popperian）的证伪检验，那么2008年的金融崩溃及随后几年世界经济的表现，已经不容置疑地证伪了这个研究纲领。没有人敢说，市场之所以在很短的时间内失去了一半的价值，是因为那些公司突然失去了一半的有用性，但公司一如既往还是那些公司。也没有人敢说，欧洲一些经济体的失业率高达20%，而且仍在上升，是因为劳动者的偏好突然完全改变了，因为人们仍然像以前一样想得到工作。2009年，《经济学人》（*The Economist*）杂志的一篇文章严厉地指出，华尔街绝不是金融危机的唯一受害者，标准的新古典经济学也是，它已经随着金融的崩溃而崩溃了。

只要稍稍反思一下，对于这种高度纯化的经济思想所遇到的困难，任何人都不应该感到惊讶。进入现代以来，尽管学习过程是非常缓慢的，但从"西方思想"学到的一个重要教训就是：如果我们努力尝试将任何一个事物还原为或化约为纯逻辑的，如我们试图确定诸如真理、存在或生命这类概念的"终极意义"，或者试图将某些研究领域，比如说将哲学、数学或数理化的经济学还原为若干狭隘的公理，那么这种努力必然

会招致失败。世界是不可能被还原为纯逻辑的，也是不可能锁入纯逻辑的铁笼里的。或迟或早，世界总有一天会突破纯逻辑，将真实的混乱一面呈现在世人的面前，上面这类研究纲领必定会以失败而告终。

新古典经济学的“纯粹秩序”正在被缓慢地取代，这个事实本身就是尊重现实的表现，这种新生的敬意是许多研究经济学的学者所共有的。稳步前进的行为经济学就是这样一种新的经济学，市场心理学则是另一种。还有，那些越来越依赖于对制度和技术的理解的经济发展理论，也是如此。本书给出的经济学新框架，即“复杂经济学”，也是如此。现在，除了我们在圣塔菲研究所的最初成员，很多人都在研究复杂经济学。

令我感到惊讶，同时也让我非常高兴的一个事实是，这种经济学新框架中的许多“现代”主题，与熊彼特、斯密、穆勒、马克思和凯恩斯等伟大思想家的思想非常契合，与许多制度主义者和政治经济学家的理论也非常吻合。他们都认为，经济涌现于技术，经济结构是不断变化的，经济不一定处于均衡状态，决策者面临根本的不确定性。复杂经济学与这些思想之间的正式联系，还没有完全建立起来。现在这种联系，更像是将我们这些新想法与过去讨论过的一些想法串联起来的思维线索。但是它们确实表明，经济学重新发现了它曾经失去的一些东西。现在，对于形成中的经济、非均衡的经济，我们终于开始有了一个统一的理论图景了。

本书中收录的文章，其写作时间跨度很大：从我1987年第一次去圣塔菲研究所时开始，一直到今天为止。不同篇章之间不可避免有一些重叠之处。其中有些文章的目的，是为了向更广泛的普通读者介绍我们的主要观点，并从不同的角度探讨这些思想。这些文章得益于许多其他学者在经济学、复杂性科学和其他领域的研究工作，特别是我在圣塔菲研究所的同事约翰·霍兰德、斯图尔特·考夫曼、戴维·莱恩和理查德·帕尔默。它们也得益于其他一些与我们当初在圣塔菲研究所的团队没有多少密切联系的学者的研究工作，特别是彼得·艾伦、罗伯特·阿克斯特尔、乔希·爱泼斯坦、阿兰·基尔曼和李·特斯法齐，他们都为这种新的经济学框架做出了贡献。此外，这些文章还借鉴了新古典主义的公式化表达方法，毕竟我所接受的学术训练就是这个。在这些文章中，有一些是全面的分析，另一些则是散文性质的文章。它们基本上是按主题，而不是按写作时间来排序的。但是从总体上看，研究文章大部分集中在本书的前半部分，而散文则集中在后半部分。读者可以按自己喜欢的顺序来阅读

它们，我鼓励读者这么做，同时也希望他们进一步阅读相关领域的文献。

如果将本书收录的所有文章合到一起来看，一个统一的思考主题或框架就会浮现出来。以往那种行为主体不复存在了，以前他们面对的是定义明确的、有明确概率结果的问题，并运用完美的演绎推理，从而可以达到均衡。而现在取而代之的是这样一种行为主体：他们必须理解他们面对的环境，必须运用手头拥有的任何一种推理方法，必须接受并做出调整以适应结果，而且他们自己的调整可能导致结果不断变化。

1996年，研究经济思想史的专家戴维·科兰德（David Colander）说了这样一个寓言：

> 一个世纪以前，经济学家站在两座高耸山峰之间的底部，而山峰则隐藏在云层当中。他们想爬上高峰，但是不得不先决定要攀爬的是哪一座山峰。他们选择了有明确定义、遵循数学秩序的那座山峰。但是，当他们费尽千辛万苦登上了那座山峰，站到了云层上之后，才发现另外一座山峰要高得多。那就是过程和有机主义之峰。

在过去这些年里，许多经济学家已经开始攀登另一座山峰了。在这个征程中发现的任何东西，我都会非常感兴趣。

COMPLEXITY

AND THE

ECONOMY

01

什么是复杂经济学

经济思想的新框架

复杂经济学以一种不同的方式来思考经济，它将经济视作不断进行自我“计算”、不断自我创建和自我更新的动态系统。它强调偶然性、不确定性、意义构建和“一切变化皆有可能”，是一门以预测、反应、创新和替代为基础的“动词”学科。复杂经济学正在取代新古典经济学而稳步走向经济学的中心。

本章是一个导论，讨论了本书其他章节的诸多主题。本章的目的是为复杂经济学提供一个框架。复杂经济学建立在经济不一定处于均衡状态这个基本命题的基础之上。经济主体，如企业、消费者、投资者，不断改变自己的行为和策略，以便对他们共同创造的结果做出反应，而且这种反应进一步改变了结果，这又需要他们重新进行调整。因此，经济主体生活在这样一个世界里，他们的信念和策略要不断地接受结果或“生

态”的检验，而这种结果或“生态”正是这些信念和策略一起创造的。

以往的经济学在很大程度上回避了如上这种非均衡观点，但是只要我们接受它，就可以观察到标准的均衡分析中无法观察到的模式或现象。这些模式或现象以一定概率出现，持续一段时间后可能会消散，并且在经济的“中观层面”上，即微观层面和宏观层面之间发挥作用。我们还认为，经济并不是给定的、一成不变地存在着，而是在一系列技术创新、制度和安排的不断发展中形成的，这种技术创新、制度和安排还会引出进一步的技术创新、制度和安排。因此，在复杂经济学看来，经济是运动着的、永远在“计算”自身的。或者说，经济无时无刻不在重新构建自身。均衡经济学强调秩序、确定性、演绎推理和静态分析，而复杂经济学则强调偶然性、不确定性、理解（即意义建构）和对变化的开放心态。

本章是专门为本书撰写的，它的基础是我于1999年发表在《科学》杂志上的论复杂性与经济的文章。当时，我正在帕洛阿尔托研究中心的智能系统实验室做外聘教授。在此，我非常感谢罗南·阿瑟（Ronan Arthur）、理查德·布朗克、戴维·科兰德、多因·法默、玛格达·丰塔纳（Magda Fontana）、奥利·彼得斯（Ole Peters）、戴维·赖斯曼（David Reisma）和威廉·塔布（William Tabb）等人的宝贵意见。

* * *

在过去的25年间，一种与以往的经济学有着全然不同的经济学观点的复杂经济学，逐渐诞生并缓慢地发展起来了。复杂经济学认为，经济不一定会处于均衡状态；在经济学中，计算与数学同样非常有用；在同一种经济状况下，既有可能出现收益递增，也有可能出现收益递减；经济不是给定的、一成不变地存在着，而是在一系列制度、安排和技术创新的不断发展中形成的。在很大程度上，复杂经济学的研究起源于20世纪80年代后期的圣塔菲研究所。不过，现在对复杂经济学进行研究的学者已经非常多了。[1]复杂经济学也引发了不少疑问。例如，复杂经济学这种不同的思考经济问题的框架能为我们带来什么，它怎样才能发挥作用，又适用于哪些领域？复杂经济学会不会取代新古典经济学，又或者它会不会融入新古典经济学？如果存在一种逻辑框架的话，复杂经济学是如何在这种框架下发挥作用的？

本章的目的就是来回答上述问题，尤其是最后一个问题的。为了达到这个目的，我在本章中要完成的工作，不是对复杂经济学进行综述，

也不是提供一份“复杂经济学导游图”，而是要给出一个思考这种新经济学观点的框架，一个连贯的逻辑框架。首先，我将论述复杂经济学的若干基本原则；然后，我将在我自己以前发表的两篇文章，以及其他一些学者研究的基础上，阐明复杂经济学的核心要点。[2]

在本章中，我将证明复杂经济学绝不仅是标准经济学的延伸，也不只是标准经济学模型再加上基于行为主体的行为。复杂经济学以一种全然不同的方式看待经济，给出了一个完全不同的图景。在这个图景中，行动和策略是不断进化的，时间变得至关重要，结构不断形成且不断重组，标准的均衡分析中不可见的现象将浮现出来并得到研究，介于微观层面和宏观层面之间的中观层面也是非常重要的。与新古典主义理论中的世界相比，复杂经济学中的世界更接近于政治经济学中的世界，那是一个有机的、进化的、充斥着历史偶然性的世界。

经济与复杂性

接下来，我们先来讨论经济本身。经济是一个庞大而又复杂的，由各种各样的制度安排和行为构成的体系。在经济这个体系中，不同的行为主体，如消费者、厂商、银行、投资者、政府机构，从事着各种各样的活动，如买卖、投机、贸易、监督、生产产品、提供服务、对公司投资、制定策略、探究、预测、竞争、学习、创新，以及调整适应，等等。用现代术语来说，经济就是一个有着无比庞大的并发行为（concurrent behavior）的并行系统（parallel systems）。市场、价格、贸易协定、制度和产业，全都形成于这些并发行为中，并最终形成了经济的总体模式或聚合模式（aggregate pattern）。

> 复杂性科学（**Complexity Science**） 一门研究系统中相互作用的要素如何生成整体模式，整体模式反过来又如何导致这些要素发生变化，或导致这些要素调整以适应整体模式的科学。

一种历史悠久的、可以追溯到亚当·斯密时代的经济学观点认为，这

种总体模式形成于个体行为，而个体行为反过来又会受到这种总体模式的影响。这里存在着一个递归循环，正是这种递归循环使经济具备了复杂性。复杂性科学研究系统中相互作用的要素如何生成整体模式，以及整体模式反过来又如何导致这些要素发生变化，或导致这些要素调整以适应整体模式。复杂性并不是一种理论，而是一种科学运动。例如，我们可以研究大量单个汽车在行驶中如何共同形成了某些交通模式，这些模式反过来又如何影响单个汽车的位置。复杂性是关于结构形成的：结构是如何形成的，这种结构形成又是如何影响和导致客体的生成的。

从复杂性的视角出发去研究经济或经济中的某些领域，就意味着要探究经济是如何进化的，也就意味着要深入详细地研究个体行为主体的行为，是如何共同形成某种结果的，以及这种结果反过来又是如何改变他们的行为的。换句话说，从复杂性的视角出发，要研究的是个体行为者如何应对其行为共同创造出来的模式，以及这种模式是如何实现自我调整的。通常来说，这是一个难度很大的问题。这就要求我们探究一个过程是如何从多个行为主体的有目的的行动中生成的。也正因为如此，历史上早期的经济学，采用了更加简单的、便于进行数学分析的方法来处理各种经济问题。这种更加简单的方法，要探究的不是行为主体如何应对他们的行为所形成的总体模式，而是什么样的行为，如行动、策略、预期是能够被这些总体模式所支持的，或者说是符合这些总体模式的，或者说是与这些总体模式相一致的。

以往的经济学要探究的是，何种模式不需要微观行为做出改变，何种模式将处于静止状态或均衡状态。比如，一般均衡理论要研究的问题是，被生产出来的商品的价格和数量，和被消费掉的商品的价格和数量如何做到一致，才会符合各个经济市场的价格和数量的总体模式，因而行为主体也就没有受到激励去做出改变。又如，经典博弈论要研究的问题是，在给定博弈对手可能选择的策略、行动和资源配置的情况下，一个行为主体根据某种判断标准，应该选择什么样的策略、行动和资源配置，才是最优的行动选择。再如，理性预期经济学要研究的是，什么样的预期符合所有这些预期共同创造的结果，或者说平均而言能够被所有这些预期共同创造的结果所验证。

这种便捷的均衡理论，不失为研究经济模式的一种自然方法，同时也为数理分析留下了用武之地。利用这种方法来推进经济学研究无疑是可以理解的，甚至可以说是适当的，而且该方法也确实带来了累累硕果。这种方法的核心理论建构是一般均衡理论。它不仅在数学上是严谨

的，而且还通过对经济的建模，使我们得以在心智上对经济进行重构，从而给了我们一种刻画经济的方法，一种全面理解经济整体的途径。这一点无疑是极其有价值的。除了一般均衡模型之外，其他理论，如厂商理论、国际贸易理论和金融市场理论的均衡模型，也都是非常有价值的。

不过，这种“均衡建模法”也不是没有代价的。许多经济学家都不认同这种均衡建模法，以及在它基础上形成的新古典经济学理论，理由是它假设了一个理想化和理性化的世界，这扭曲了现实世界，而且为这个世界所选定的基本假设也往往只是为了便于分析。[3]我本人也持同样的反对意见。正如许多其他经济学家一样，我也赞赏新古典经济学的美好，但在我看来，新古典经济学的理论建构过于纯粹、过于脆弱了，必定会在现实世界面前碰个头破血流。新古典经济学生活在一个有序、静态、可知、完美的柏拉图式的世界当中。在这个世界里，现实世界的模糊不清、混乱无序和真实多变是完全不存在的。

当然，所有杰出的经济学家都非常清楚，现实世界中的经济远比新古典经济学世界中的经济更加丰富多彩。有人或许会认为，我们可以这样做，一方面坚守均衡概念作为我们思考经济问题的基础，同时在另一方面，通过经验和直觉来填补更加丰富的现实内容。但是，这还远远无法令人满意。如果我们预先假定经济是均衡的，那么就等于设置了一个过滤器，我们就无法观察到经济中的很多现象。这是因为均衡本身的定义就决定了，均衡的经济没有改进或进一步调整的余地，没有探索和创造的空间，也没有任何暂时性现象的存身之地。在均衡理论中，任何可以在经济中带来改变的东西，如适应、创新、结构变化及历史本身，都会被绕开或被忽略。由此而产生的新古典经济学，也许算得上一个完美的理论建构，但是它缺少真实性、活力和创造力。

那么，如果“允许”经济学考虑更多的可能性，并且能够研究行为主体如何应对他们共同创造的模式，又会怎样呢？这会不会使一切都变得有所不同呢？到时候我们又将会有什么样的发现呢？

内生的非均衡

首先要指出的一点是，只要我们探究“行为主体可能如何应对”这样一个问题，其实就已经隐含地假定经济是非均衡的了。这是因为，如果新的应对方法是可能的，那么这些方法就会改变结果，所以由定义可

知，这不可能是均衡的。接受过良好的新古典经济学训练的经济学家，可能会对非均衡假设持反对意见，因为标准经济学理论认为非均衡情况在经济中无足轻重。萨缪尔森在1983年就曾经说过："那么不稳定的非均衡状态，即使真的存在，也必定只是暂时的、非持久的状态……读者诸君，你们几时曾见过竖起来的鸡蛋呢？"[4]

这些经济学家明确地告诉我们：均衡就是经济的自然状态。

但是，我认为事实并非如此，绝对不会如此。我敢肯定，非均衡状态才是经济的自然状态，因此经济始终处于变化当中。这不仅是因为经济总是面临着外部冲击或外界影响，而且还因为非均衡本身就产生于经济的内部。内生的非均衡的出现，主要出于两个原因，第一个原因是根本的不确定性或奈特意义上的不确定性，第二个原因是技术创新或技术变革。下面让我们依次来阐述。

第一个原因是根本的不确定性。经济中所有关于选择的问题都与未来发生的事情相关，这些事情既可能马上就会发生，也可能过段时间才会发生。因此，经济中的选择问题，必定在一定程度上与未知的事情相关。在某些情况下，行为主体拥有充分的信息，或者能获取可能会发生的事情的准确概率分布，但是在很多其他情况下，即在绝大多数情况下，他们并不拥有这些信息，甚至根本不知道这些信息，也无法估计出概率分布。[5]例如，我可能会选择某一种新技术进行风险投资。在刚开始时，我可能完全不知道这种技术会不会成功、公众对它的接受程度如何、政府会对它进行怎样的监管，也根本不知道会不会有人把同类产品引入这个市场。然而，我必须有所行动，因为我对相关的情况，即根本的不确定性完全一无所知，所以所谓"最优"行动根本就不存在。当其他行为主体也参与进来时，情况会变得更糟。在那种情况下，这种不确定性就会自我强化。如果我不清楚具体情况，我只能认为其他人也不清楚。我不仅必须形成自己的主观信念，还必须形成有关主观信念的主观信念，而且其他行为主体也必须这样做。由此，不确定性带来了更进一步的不确定性。[6]

当然，我的这个观点并不是一个新观点。其他经济学家，尤其像沙克尔（Shackle），已经撰写了大量论著论述这一点。只有将这个观点理论化，才能显示出重要的理论意义。当我们无从得知结果是什么时，它所带来的决策问题是无法清晰明确地加以界定的。问题本身都未能在逻辑上界定清楚，针对这些问题的原由也就不可能有一个合乎"逻辑"的解决方法，因此理性，即纯粹的演绎理性（deductive rationality），也不

可能得以明确地界定。在这种情况下，演绎理性不仅只是一个糟糕的假设，而且它本身根本就不可能存在。也许有可能存在聪明的行为、合乎情理的行为，也许可以存在富有远见的行为，但是从严格意义上来说，根本不存在合乎演绎理性的行为。因此，我们不能假设这种理性。

当然，所有这些并不意味着人们面对经济问题时束手无策、寸步难行，也不意味着人们不会做出选择、采取行动。行为经济学告诉我们，情境往往决定了人们如何决策，也无疑让我们直接利用它来“替我们发现”结果。此外，认知科学也告诉我们，如果某个决策很重要，那么人们有可能会摆脱当前情境的影响，他们会努力通过推断、猜测，以及利用过去的知识和经验去分析这个决策。他们会发挥自己的想象力，尽可能地预测未来，并以此为依据来做出决策。确实，正如沙克尔所指出的，“每个人都会发挥自己的想象力，努力想象未来的样子，这种想象过程是其决策过程中一个至关重要的组成部分。”依据沙克尔等人的这种洞见建立决策模型的一种方法是：假设行为主体对自己身处的环境形成了某种个人信念或假设，也有可能是一系列信念或假设，即内部模型，并且不断地对他们的信念或假设进行更新。这也就是说，当他们在探索时，他们以此为依据不断对自己的行动和策略进行调整、舍弃和替换。[7]总之，他们是在利用归纳不断前行。[8]

这种探索行为的不断实现，导致了经济中永恒存在的布朗运动。由于行为主体的探索、学习及适应，经济永远都处于破坏性运动之中。正如我们将会看到的那样，这些破坏性运动会被放大，成为显著的现象。

破坏性运动的另一个动力是技术变革。大约在100年前的1912年，熊彼特提出了一个著名的观点，他指出“经济体系中存在着一种力量，这种力量能够破坏任何可能达至的均衡”。这种力量来源于“生产方式的新组合”，我们现在称之为技术的新组合。经济学并没有否认这一观点，但是它必须假设经济均衡能够不时地进行调整，以适应外界的变化。

然而，这种技术力量的破坏性很强，远远超出了熊彼特所设想的范围。新技术需要更多的其他新技术来支持。例如，在人们发明了电脑以后，电脑就需要或“要求”更强大的数据储存技术、计算机编程语言、计算算法及固态开关设备，等等。而且，新技术也为其他新技术的出现创造了条件。举例来说，真空管的问世，使得无线电的传送和接收、广播、继电器电路、早期计算机以及雷达等技术的出现成为可能，或者说它为后来的这些技术提供了“供给”。同时，这些新技术反过来又催生了对更新技术的需求和供给。由此可见，一项新技术并不是只会使均衡受

到一次性的破坏，相反，新技术永远都是更新技术的创造者和需求者，而且这些更新技术本身，也需要创造出比自己更新的技术。我要再次提请读者注意这个过程自我强化的性质。由此而导致的结果并不是偶发性的破坏，而是持续性的、一浪催生一浪的破坏大潮。在整个经济中，这种破坏并行出现，在所有维度上同时发生。技术变化会内生地、不断地创造出更进一步的变化，从而使经济处于永远的变化之中。

从时间维度来看，技术的破坏性影响发挥作用的速度，要慢于纯粹源于不确定性的布朗运动。但是，技术的破坏性影响会导致更大的剧变。技术的破坏性影响本身，也会带来进一步的不确定性，因为各行各业，如工商界和产业界，根本不可能知道下一步进入自己领域的技术会是什么。因此，不确定性和技术这两个因素，都会导致这样一个现象：行为主体没有任何确定性的方法来做出决策。

现在，一种全新的看待经济的方法正在崛起，它不同于标准的均衡性经济学方法。由于在经济中，不确定性和技术变革无所不在，而且毫无疑问，它们二者渗透了经济的各个层面，行为主体必须探索着前行、“学习”自己面临的决策问题并对出现在眼前的机会做出反应。在我们所处的世界里，行为主体的信念、策略和行动创造了某种状态、结果或“生态”。而与此同时，行为主体的信念、策略和行动，都要接受这些状态、结果或生态的“检验”。再者，更加微妙的是，行为主体进行的探索，还会进一步改变经济本身以及行为主体所面对的环境。因此，行为主体不仅要面对自己试图解决的问题，而且他们在解决问题过程中的每一个行动，合起来还会改变当前的结果，这就要求他们必须再次做出调整。换句话说，我们处在一个复杂的世界当中，这种复杂性与非均衡有着密切的联系。

建立非均衡状态下的理论

面对内生的非均衡，我们应该怎么办呢？如果庞大的经济一直随着行为主体的活动处于“沸腾”之中，借用熊彼特的一句话，那么我们要处理的似乎就是一种“无法纳入分析范围的混沌”。面对这个难题，以往的标准经济学的态度可以用两个成语来描述：束手无策和退避三舍。但是，如果我们决定不步标准经济学的后尘，并且我们坚定自己的立场，认真对待非均衡问题，那么我们必须怎样做才能继续向前呢？我们能够得出一些有用的结果吗？我们会有什么样的发现呢？当然，首先要回答

的一个问题是，在非均衡状态下进行理论化建模意味着什么。

有一种观点认为，经济的很多组成部分可以被视为处于近乎均衡的状态，对于它们，标准理论仍然是适用的。同时，对于经济的其他组成部分，则可以将它们视为处于暂时偏离了最具吸引力状态的状态，仍然可以研究它们向这个最具吸引力的状态收敛。但是，这种观点仍然把经济当成了一台高度平衡、能够自动调整的机器，认为它只会暂时偏离均衡状态。固守这种观点，只会使我们既不能了解经济在均衡状态之外的表现，也不能刻画经济在非均衡状态下极具创造力的一面。

研究非均衡经济的一种更好的方法是，研究经济的各种“当前状况”。正是在这些当前状况中，形成了决定未来事件或事物的那些条件。经济是一个系统，而且这个系统中的各个元素，都会根据“当前状况”来不断更新自己的行为。[9]如果采用另外一种更加正式的说法，我们可以说，经济就是一种持续的“计算”（computation）。这是一种极其庞大的分布式计算，也是一种大规模的并行计算，而且这种计算是随机的。[10]这样一来，经济就可以视为一个以一系列事件为序不断进化的系统。从这个角度来看，经济是有算法规则的。

虽然以这种方式看待经济有一个风险，有人可能会说，这只是为了迎合科学的当代潮流。但是基于这种思想，我可以阐明一个很重要的观点。让我们暂且假设我们掌握了经济的算法，或者更进一步地说，假设自己就是拉普拉斯（Laplace）或“上帝”[11]那样的人，在我们掌握的经济或感兴趣的某个经济领域中，“采取”下一步行动时，所要遵循的数量庞大但总数有限的各种具体机制。有关计算的一个基本定理告诉我们，一般来说，如果我们随机选择了某种算法，是没有方法，或者说没有系统的解析方法能够提前算出该算法或电脑程序是否会终止，而不是永远持续或循环下去的。因为我们只能规定，如果某种算法的输出，符合一组特定的数学条件或得到了某个给定的“解”（solution）就终止计算，那么一般来说，我们并不能确定这种算法是否合适。总之，没有任何解析方法能够提前确定某种给定的算法是否合适。[12]我们所能做的，无非是按照算法计算下去，然后看看它会带来什么结果。如果某个算法足够简单，我们还是经常可以观察到，它会带来某种特定的结果。但是，当我们不能决定算法的结果时，算法就不必过于复杂了。

非均衡系统
（**Nonequilibrium**

System） 一些经济学家认为，非均衡状态才是经济的自然状态，经济始终处于变化当中。这不仅是因为经济总是面临着外部冲击或外界影响，而且还因为非均衡本身就是产生于经济内部的。

因此，我们必须更加谨慎一些。对于一个高度互联的体系来说，均衡或闭合解（closed-form solution）都不是缺省结果。而且，如果均衡或闭合解是确实存在的，那么必须解释它们存在的理由。从计算的角度思考这些系统，并不意味着我们有意回避解析分析，严格地说，解析分析是非常必要的。我们经常要对非均衡系统的定性特点进行很多非常有用的预分析，以便更好地理解它们背后的机制。然而，在研究非均衡系统的结果时，唯一准确的方法仍然是计算。

现实经济背后的算法并不是随机选择的，而是高度结构化的。因此，一种可能出现的情况是，现实经济的“计算”总是会得到非常简单的结果；另一种同样可能出现的情况是，现实经济的计算也总是无序的、无定形的。在我们所研究的经济领域内，通常不会出现这两种情况。尤其是在有强大的抗衡力量发挥作用的情况下，我们经常可以观察到一些大型结构，即一些与均衡不严格对应的吸引域。在这些吸引域内或当不存在吸引域时，我们也能观察到某些机制，它们会造成某些不是随机产生或消亡的现象、子模式或子结构。对此，我们可以用物理学中研究的太阳来进行类比。从远处看，太阳是一个由气体组成的巨大球体，而且是一个处于均衡状态之下的球体。但是在这个“均衡”的内部，还存在着一些强大的机制，它们引起了许多动态现象，如巨大的磁环和磁拱、冕洞、X射线耀斑，以及最高时速可达7.2×10^6千米的等离子射线大规模爆发等。太阳这个巨型“气球”确实呈现为一个松散的球体，但是它从来都没有处于均衡状态。相反，它一直处于不断的运动之中，这种运动源于更早之前的扰动，而且它破坏了达到均衡的可能性。这些现象都是局部的，并且能够发生在各种维度上。再者，这些现象都是短暂的，它们的出现、消失和互动，从时间上看都是相当随机的。

我们在经济中也经常可以观察到类似的情况。要建立非均衡状态的理论，就是要揭示那些起作用的大吸引子（如果它们真的存在的话），

同时还要研究其他子结构或现象，这些子结构或现象可能因大吸引子的特点和行为而出现。我们可以利用精心设计的计算机实验来做到这一点，通常是对结果进行统计分析，从而将各种现象及导致这些现象的机制识别出来。在很多情况下，我们可以为某种现象建立一个较简单的“玩具模型”（toy model），该模型应该能够刻画该现象的基本特征，并允许我们利用数学理论或随机理论来研究这种现象。但是要记住，研究的目标并不一定是要给出确定的方程式或达到某些必要条件，相反，正如所有的理论一样，我们的目标是获得一般性观点。

接下来，让我们通过一个真实的、利用计算机完成的非均衡研究，来将上述要点融合起来。这是一个经典案例。

1991年，克里斯蒂安·林格伦（Kristian Lindgren）设计了一个在计算机上进行的锦标赛。在这个锦标赛中，各种策略随机配对，进行重复的囚徒困境博弈，以便分出高下。在这里，我们不必考虑囚徒困境博弈的细节，而是直接把它视为一个有一系列指定策略的简单博弈。所谓博弈策略，就是指给定对手最近采取的行动，另一方应该如何行动。如果某个策略带来的结果很好，那么就重复该策略并进行策略突变；如果某个策略带来的结果很糟糕，那么该策略就会被移除。林格伦允许博弈参与者拥有对另一方和自身最近采取的行动的深层记忆，从而可以“深化”策略。这样一来，用我们在这里采用的术语，就可以说这些策略在“探索”策略空间。如果策略不是很成功，那么就可以进行改变和调整。林格伦发现，在锦标赛开始之初，简单策略，如“一报还一报”策略是占优策略，但是过了一段时间后，“更深层”策略出现并战胜了原来的简单策略。随着时间的推移，又出现了能够“剥削”以前更深层策略的更加深层的策略，这个过程是在相对稳定期间和动态不稳定期间的相互交替中完成的（如图1-1所示）。

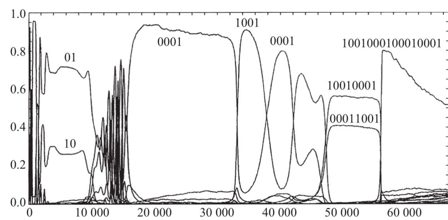


图**1-1**　林格伦计算机锦标赛中的策略

图中横轴表示时间（期数），纵轴表示使用某个特定策略的次数，标号表示策略的记忆深度。

这个锦标赛的动力学机制十分简单，因此林格伦可以将它们用一些

随机方程式描述出来。但是，这些随机方程式不能说明全部情况，我们必须通过计算来搞清楚到底会发生什么。在计算过程中，我们发现涌现出来的是一个生态，即一个“策略生态”。每种策略都试图利用某个环境，在该环境下求得生存，而且该环境就是由该策略本身以及其他策略在努力利用环境、寻得生存时所创造的。这个生态就是一个微型的“生物圈”，在这个生物圈中，各种新物种（即策略）不断涌现出来，在现有各物种所创造的环境中探索求生，如果遭到失败，这些失败的策略就无法生存。这里需要提请读者注意的是，这个生物圈中当然也有进化，但这种进化并不是从外部引入的，而是在各种策略为生存而竞争的自然趋势中发展出来的。这种观点在复杂经济学这种类型的经济学中是很常见的。复杂经济学中的“解”，是一个由相互竞争的多种策略、行为或信念组成的生态系统。这个生态系统是不断变化的，它拥有自己的特性，对它可以进行定性研究和统计研究。[13]

在林格伦的这项研究中，每一轮计算的结果都各不相同。不过，在经过多轮计算之后，终于出现了一个进化稳定策略，那是一个复杂的策略，它依赖于对过去四期行为的记忆。而且，在其他各轮锦标赛中，这个系统仍然持续不断进化。在某些轮次中，我们观察到复杂的策略很快就出现了，而在另外一些轮次中，复杂策略则很迟才出现。尽管如此，这个锦标赛中还是存在一些不变的东西，如策略之间的共存现象、新策略的开发、自发涌现的互利主义、忽然发生的崩溃、静止状态和不稳定状态之间的交替变化，等等。这些情况与古动物学上的图景何其相似！

我在这里将林格伦这项研究称为非均衡经济学研究的一个样板。有的读者可能会心生疑虑：对一个在计算机上进行的研究，怎么可以算是经济学研究呢？这种研究与建立非均衡状态下的经济学理论，有什么关系呢？这看上去一点也不“数学”。对于这种疑问，我的回答是理论绝非全由数学构成。数学无非是一种技术、一个工具而已，尽管它看上去比较精确、比较复杂。理论不同于数学，理论就在于发现、理解并解释世界中存在的现象。数学只是为这个理论化过程提供便利，当然这是一个很大的便利。重要的是，计算也能起到同样的作用。

当然，计算与数学也有不同。利用数学模型时，我们可以通过方程一步一步地论证，并找到问题的解必须满足的条件，计算却不能做到这些。[14]计算也有自己的长处。它的长处不但能补偿它的不足，还可以让我们看到均衡数学无法看到的现象。通过计算，我们能在不同的条件下重新得到结果，在结构出现或没有出现时进行探索，确定潜在的深层机

制，层层递进地简化现象，提取现象的根本信息。换句话说，计算是思想的助手，在这一点上，它与经济学早期发展中所运用的其他辅助工具没有什么区别。线性代数、微积分、统计学、拓扑学、随机过程等辅助工具，在当时都曾经受到过抵制。计算机已经成了研究经济学的一个实验室，如果能够熟练地、有效地利用计算机，它可以成为一个强有力的理论创造器。[15]

所有这些都指向一个新的前进方向，即以非均衡视角来研究经济的方向。我们可以将经济或我们感兴趣的部分经济领域视为行为主体的策略、预测和行为不断变化的结果。对于这些经济领域，以及经济学中的一些经典问题，如代际转移支付、资产定价、国际贸易、金融交易、银行业务等，我们都可以通过建立模型来研究。只不过在我们的模型中，要研究的不只是行为主体在均衡状态下做出的应对，而是行为主体在所有情况下做出的应对。我们的模型有时也可以借助于数学来进行分析，但是许多时候只能借助于计算，当然有时需要同时借助于这两者。我们不仅希望找到均衡的条件，我们还想理解结果的形成以及结果的进一步发展，解释经济中出现的所有动态现象。

三种典型的非均衡现象

那么，在非均衡状态下会出现什么现象呢？这些现象同非均衡与复杂性有什么关系？接下来，我将依次对这些问题展开分析。在研究非均衡经济中可能出现什么模式和结构之前，我们不妨先来看一下，均衡这个过滤网下的模式与我们所见的模式之间有什么不同之处。为了说明我们的观点，下面先考虑一个简单的交通流量模型，尽管它与经济的关系似乎不是那么紧密。

一个典型的交通流量模型通常都会承认这一点：当一辆车与前面的车辆离得很近时，该车应该减速；与前面车辆相距很远时，该车应该加速。如果给定某个较高的交通密度，例如每千米有N辆车，那么就意味着车辆之间有一定的平均间距，车辆应该放慢车速或提高车速，以便与之相适应。这样一来，也就在不经意间出现了一个均衡速度，如果我们希望得到的解只限于均衡状态，那么我们就只能看到这种模式。在现实生活中，当交通密度较高时，往往会出现非均衡现象。有些车辆会放慢速度，因为司机可能注意力不够集中或受到了干扰，而这就会导致它们后面的车辆随之减速，从而立即压缩了交通流量，并进而导致后面的车

辆进一步减速。随着这种压缩不断向后蔓延，交通就会受阻，交通堵塞就出现了。然后，等过了一段时间之后，交通又会恢复正常。这里需要注意三点：第一，这种现象是自发的，每次出现的时间、蔓延的长度以及恢复的时间都是各不相同的。这也就意味着很难找到闭合解，因此最好利用概率方法或统计方法来对此进行研究。第二，这种现象是暂时的，是在一定时间内出现或发生的，如果我们坚持均衡观点就不会观察到这种现象。[16]第三，这种现象既不是发生在微观层面上的（即单个车辆），也不是发生在宏观层面上的（即道路上的所有车辆），而是发生在这两个层面之间，即中观层面上的。

那么，更一般的非均衡经济又会是怎样一种情况呢？如果将均衡这个过滤网移走，我们会看到什么现象呢？这些现象又是怎样发生和发展的呢？接下来，作为例子，我将讨论三种非均衡现象。

第一，资产价格变动的自我强化，或者用通俗的说法就是“泡沫和崩溃”。只要看一下圣塔菲研究所的人工股票市场，就可以了解这种现象是怎么产生的。在这个基于计算机的模型中，“投资者”是一些人工智能计算机程序。根据我们在前面已经阐述过的理由，这些“投资者”不能简单地假设或推断出某个给定的“理性”预测模型，相反，他们必须分别去发现某个有效的预期模型或预测模型。这些“投资者”会随机地创造出或发现他们自己的预测模型，试用那些有“应用前景”的预测模型，舍弃那些没有用的预测模型，而且他们还会定期地创造出新模型来替换旧模型。股票价格在这些“投资者”的买卖过程中形成，因此最终形成了行为主体的预测。这样一来，我们这个市场成了一个预测模型的生态系统，这些预测模型要么成功，要么被淘汰出局，该生态系统因而处于不断变化当中。[17]在这个人工模型股票市场中，我们可以观察到很多现象，而其中最主要的就是自发形成的泡沫和崩溃现象。

要想搞清楚这些现象是怎样产生的，我们可以从这个实验中提取一个简单的机制。假设我们的“投资者”之中，有人发现了如下交易预测规则：“如果股票价格在最近的k个交易期内上涨，那么就预期价格会在下个交易期内上涨$x\%$。”同时假设，有的“投资者”，也有可能就是上述这些“投资者”，发现了如下这样的预测规则：“如果当前的股票价格是基础盈利或股息的y倍，那么就预期价格会下跌$z\%$。”第一种预测可能会导致泡沫行为：如果价格上涨了一段时间，“投资者”就会买进，这样也就证实了这种预测，从而就可能导致价格的进一步上涨。到最后，当这种预测驱动价格上涨到一定高度后，就会引发第二种预测。于是，持有这些

股票的“投资者”会抛售这些股票，股价下跌，这样就会终止上涨的预测，也导致其他“投资者”跟着抛售股票，最终就会致使股票崩盘。这种扰动的规模和持续时间各不相同，而且发生的时间也很随机，因此是不可预测的。唯一可以预测的是，这种扰动的现象肯定会发生，并且振动的规模大小有一定的概率分布。

第二个暂时现象是集群波动（**clusted volatility**）。所谓集群波动，是指低波动期与高波动期随机交替出现的现象。在我们的人工股票市场中，集群波动表现为价格低波动周期与高波动周期的交替出现。当行为主体的预测规则在相当程度上相互一致且能够起作用时，就会出现价格低波动周期，这时行为主体没有什么动力去改变这些预测规则或这些预测所产生的结果。当一些行为主体发现了更好的预测规则（即“预测器”）时，就会出现价格高波动周期。因为这会打破整体的模式，使得其他“投资者”不得不改变他们的预测规则来重新适应环境，这就会导致进一步的扰动，以及进一步的重新适应新环境。这种模式在林格伦的研究中，可以看得非常清楚（见图1-1）。由此而导致的结果是，在一段时间内，会出现频繁的再调整或激烈的波动。在现实的金融市场数据中，这种随机的低波动期和高波动期交替出现的现象，被称为“广义自回归条件异方差行为”，即GARCH行为。

第三个现象是被我们称为突然渗透（**sudden percolation**）的现象。这种现象更经常地发生在空间的维度上，而较少发生在时间的维度上。在一个网络内，当某个地方出现了可以传播的变化时，如果这个网络内部的联系比较“稀疏”，那么这个变化就迟早会因为可用的“转接”不足而逐渐消失。如果网络内部的联系很紧密，这个变化将会不断地传播下去。在银行网络中，某个银行可能发现自己持有不良资产，于是该银行就有压力去提高资产的流动性，并向作为其交易对手的银行求助。这样一来，作为该银行交易对手的那些银行也会面临压力，不得不提高自己的资产流动性，于是又向它们交易对手的银行求助。因此，不良资产问题很快就会通过“多米诺效应”传遍整个银行网络。这样的事件会造成很严重的破坏。这种问题在一些联系不紧密的网络中会逐渐消失，但是如果网络内部的互联程度超过了某个阈值，并且联系变得更加紧密之后，这些问题就会在很长时间内持续地传播，甚至渗透到整个网络。[18]

上面的最后一个例子，让我们了解到了复杂系统的一个一般性质。通常，在复杂系统内，只有当模型中所刻画的调节强度或联系程度的基本参数值超过某个阈值或达到了某个临界水平时，一些现象才会出现。

在这个关节点上，系统的整体行为会出现一种相变（phase transition）。在我们的人工股票市场中，“投资者”以一个较慢的速度搜寻新的预测规则，市场行为会“坍塌”为理性预期均衡。不同行为主体会做出同样的预测，而这些预测会使价格发生变化，价格变化的总体情况通常又会证实这些预测。在这种情况下，简单行为占了主导地位。但是，当我们的“投资者”搜寻新预测规则的速度变得非常快且更加符合实际时，市场就会形成一种“复杂心理”，产生各种不同的预测信念。这时，各种各样的暂时现象就开始出现了。在这种情况下，复杂行为就占了主导地位。当我们继续调高“投资者”搜寻新的预测规则的速度时，个体行为就不能有效地进行调整来适应他人行为的快速改变，于是混沌行为就会占据主导地位。其他的一些研究也发现了从均衡到复杂再到混沌的相变，或者从均衡到复杂再到多重均衡的相变。我认为，在非均衡模型中，一般都存在着这种相变。

现在，我们可以开始了解这种现象（如果你愿意的话，也可称其为秩序或结构）与复杂性之间有什么联系了。正如我所指出的，复杂性科学研究的是相互作用所产生的结果，它研究各种元素，如粒子、细胞、偶极、行为主体、企业等的相互作用中产生的模式、结构或现象。很明显，这种相互作用同样发生在我们上面说的这个网络案例中，不过在我们的人工股票市场中，相互作用显得更加微妙一些。只要我们的“投资者”中有一个人买入或抛售股票，这种行为就会导致股票价格发生变化，虽然也许只是非常微小的变化，但其他投资者就可能会对这种变化做出反应。在上面提到的所有三个例子中，变化都会在系统中扩散出去。

复杂性科学研究这种变化是如何“进行到底”的。或者换一种说法，复杂性科学研究的就是，这种变化是如何通过相互联系的行为扩散出去的。在银行体系中，一个银行在面对压力时，可能会将这种变化转移给与自己有联系的同伴，而这些同伴又可能将其传递给它们自己的同伴，那些同伴的同伴又可能进一步传递给它们自己的同伴……因此，在某一个节点发生的事件，可能会引发一连串级联放大的事件。这种级联事件或连锁反应，通常只会进一步对其他一两个因素产生影响，有时也会对更多的因素产生影响，只有在极少数情况下才会对很多因素产生影响。这个过程中的数学理论，是复杂性理论的一个非常重要的组成部分。该理论表明，这种事件会进一步引发其他事件，它们的传播有一些典型特征，如幂次法则（幂律，由很多小型且频繁的传播引起，只有极少数由大型且罕见的传播引起）、重尾概率分布（长程传播虽然罕见，但是仍然比正态分布所预测的更加频繁）[19]，以及长程相关性（事件可以长距

离、长时间传播）。事实上，所有系统，包括物理系统、化学系统、生物系统、地理系统等都有这种特征，即事件可以在系统中传播。在我们上面所举的与经济相关的例子中，传播也发挥着非常重要的作用，并且也具备这些特征，这其实不足为奇。[20]在现实的经济数据中，这些特征也都是显而易见的。

除了这些特征之外，我们还可以观察到其他一些东西。如果从外部改变一个系统内部各因素相互作用的程度，如提高某些事件进一步引发其他事件的概率，或者增加系统内部的连接数等，系统就会受到影响。如果原本存在某种后果的话，这种后果会从轻微影响发展到严重影响，再从严重影响发展到永久影响。系统会经历一个相变。所有这些特征都是复杂性的标志。

我们终于可以说清楚，为什么非均衡与复杂性是相互联系的。经济中的非均衡现象，迫使我们去研究非均衡导致的各种变化的传播，而复杂性科学要研究的在很大程度上就是这种传播。由此可见，非均衡经济学可以适当地纳入复杂性研究的范围之内。[21]

在这里，还需要对上述观点稍做进一步的说明。我在前面解释过的那些现象，经常会先出现在特定的历史时间或空间上，然后又消失。如果我们坚持均衡的立场，那么就无法观察到这些现象。这些现象不是局部现象，它们出现在局部网络或股票市场的某个组成中，并可能向外扩散。通常来说，它们会在各种各样的维度上发生。网络中的有些事件可能仅仅涉及很少的几个节点，而有些事件则可能会涉及整个经济系统。这些现象通常都是介于微观和宏观之间，因此我们可以恰当地称之为中观现象。[22]它们具有中观经济（meso-economy）的特征。

有人也许仍然会认为，这种现象是无关紧要的。毕竟，这样的系统中还隐藏着标准的均衡解，而且均衡解的有效性是最高的。对我们的股票市场模型来说，也许确实如此，因为没有任何一个股票市场能长期保持100倍的市盈率。[23]但是，这是一个至关重要的“但是”，正是由于一些暂时现象的存在，市场中才会发生一些有趣的事情，而且这些事情都发生在偏离均衡的时候。说到底，只有这种时候，才是能够赚到钱的时候。对此，我们可以用以下例子来进行类比说明。由于总是存在着重力，在地球上没有物体能“摆脱”重力，海洋中存在着一个近乎均衡的海平面，这个海平面的有效性最高。这一点当然是千真万确的。正如在股票市场中一样，在茫茫大海中，有趣的事情通常不会发生在均衡的海平面上，而且这种均衡的海平面很少见。有趣的事情通常只会发生在那些

永远都波动不休的海面上，而且这种波动还会造成更进一步的波动，那才是船只停留或航行的地方。

在这一节中，我用了三个相当著名的例子来说明何为复杂现象以及它们是怎样出现的。我们也注意到了其他的一些现象，当然还有更多的现象有待于我们去发现。这些现象到底是什么，它们有什么特性，以及现象之间的相互作用如何，这些都是未来我们要研究的重要问题。但现在最重要的是，上面的论述表明，我们需要关注经济中一个“新”的层面，即中观经济。在中观经济中，事件能够在各种维度上进一步引发其他事件。经济中存在一个中间层面或中观层面。正是在这个中观经济层上，各种现象出现了，这些现象会持续一段时间，然后消失不见。

正反馈

上面我们讨论了复杂经济学的机制，关于这种机制我还想进一步说明一点。这些机制源于相互作用中自我强化的行为。比如，行为主体买进了某只股票，或者轻微地扰动了市场，又或者传播了某种变化，就有可能导致进一步的买入行为，或者进一步的扰动，或者进一步的传播。又如，我们在前面已经看到的，行为主体在做出选择时往往表现出不确定性，这会导致进一步的不确定性；或者行为主体引入一些新技术，这又往往会带来更新的技术。这种正反馈会打破现状，导致不均衡，而且也会导致某种结构的出现。在交通运行中，一个小小的阻塞会引起更进一步的阻塞，进而导致一种结构的形成——交通堵塞。这也正是我所说的布朗运动出现的地方。它会导致一些小小的摄动，围绕着这种小小的摄动，会形成小型的“成核运动”，而正反馈则会放大这些运动，然后这些运动会被“锁定”，最后，随着时间的推移，它们最终会消失。

正反馈可以说是复杂系统的定义特征，或者更确切地说，正反馈和负反馈同时存在、共同作用，是复杂系统的定义特征。如果一个系统只存在负反馈（在经济学中，这就是收益递减），那么系统很快就会收敛到均衡状态，表现出“死的”行为。如果一个系统只存在正反馈，那么系统会偏离均衡，表现出爆炸性行为。只有在同时包含正反馈和负反馈时，系统才会表现出“有趣的”或“复杂的”行为。在正反馈的作用下，各种相互作用会相互叠加，形成某种结构，在经过一段时间后，又会被负反馈作用抵消，最后消失。因此，结构形成，然后又消失，其中有一些结构还会进一步继续发展，或者导致进一步的结构形成。这样的系统才

是一个“活的”系统。

这些观察结果，丰富了经济学中关于正反馈或收益递增的早期文献。举例来说，如果一个企业在产品、技术、地理区域等方面处于领先地位（这可能是由某些小概率事件所导致），那么在收益递增的情况下，它就会拥有更进一步的优势，从而可以获得更加领先的地位，那么接下来它很可能继续主导竞争结果。如果有N家企业相互竞争的话，那么最少会有N种结果，但是N的值不一定很小。在19世纪末期，打字机键盘的布局有很多种，各种不同的键盘布局为了让更多人使用而相互“竞争”，最终只有我们现在正在使用的这种布局胜出，成为了标准的打字机键盘布局。但是，一个简单的计算表明，当初其实存在着多于10^{54}种可能，无论怎么看，这都可以说是一个天文数字。

收益递增的形成过程现在已经众所周知了。在这里，我要补充的是，经济中的正反馈的普遍性，远远超出了我们以往的想象。它不仅存在于企业或产品中，也存在于各种小型机制和大型机制中，还存在于决策行为、市场行为、金融行为及网络动力学中。正反馈在所有维度上都能发挥作用，它使经济变得不稳定，甚至在宏观层面上也不例外。例如，凯恩斯理论所描述的机制，可视为一种正反馈。这种正反馈被暂时锁定在充分就业和失业这两种可能状态中的一种上。此外，正反馈还会给复杂系统带来一系列独特的属性，如多吸引子、不可预测性、无效，以及路径依赖等。在物理学中，则相对应地表现为多重亚稳定状态、不可预测性、相锁定（或模式锁定）、高能基态，以及非遍历性等。不过，我们还是可以把这些属性与正式的复杂性联系到一起。

经济的形成

接下来，我想讨论一个非常不一样的主题。这个主题建立在前面提到的技术的破坏性影响的基础之上。到目前为止，我们已经看到，给定组成经济的要素，这些要素会不断地对它们自己创造的模式做出反应，并且不断地形成不同的模式。但是，这仍然不足以刻画经济的基本特征。说到底，经济是不断地通过创造新要素来实现自我创造和再自我创造的。这些新要素通常是新技术和新制度，随着经济的进化，它们会产生新的结构。那么，这究竟是如何发生的？经济是如何形成其自身的？经济又是怎样发生结构性改变的？熊彼特在1980年将上述问题中要解释的现象，称为“我们试图解释的现象中最重要的一个”。复杂性思想有助

于解释这个问题，因为复杂性就是关于结构的创造和再创造的。

我们首先要认识到：如果想搞清楚经济是如何建构自身以及如何改变的，那么我们就必须先去研究技术，搞清楚技术本身是如何建构自身以及如何随时间改变的。当然，技术并不是造成经济改变的唯一因素，但很显然是最主要的一个因素。关于经济变迁的标准，经济学理论将技术与生产函数等同起来，并把经济视为这些技术的“容器”。当新的工业技术被引入进来时，生产函数就会发生变化，于是产量也提高了，劳动力或其他资源得到了释放，而这就意味着可以投入更多的财富，用来研发更新的技术。从而，经济平滑地从一个均衡转移到另一个均衡，并实现了内生增长。这种模型很漂亮，也非常符合均衡经济学的套路。但是问题在于，它使经济的主要驱动力——技术，成了背景因素，而把价格和数量放在了前台显著位置。这种观点把技术看作是无形无相的，技术自然会神不知鬼不觉地翩然而至，一个一个地随机出现，而且没有结构可以用来解释技术是如何产生的，又如何随时间的推移而改变经济。

根据复杂经济学的观点，技术是处于最显著的前台位置，价格和数量反而处于背景位置。[24]从复杂经济学视角出发，我们将会认识到，经济中的重要结构是可以用于解释技术的兴起以及技术是如何进入经济的。为了得到这种认识，复杂经济学聚焦于任何一个时点上出现的技术集合（collection of technologies），并且追问这个技术集合是如何进化的，即集合中的技术是怎样产生的，这个技术集合又是如何创造和再创造一个相互支持的支撑体系（supporting set）的，这个支撑体系又是如何随时间的推移改变了经济的结构。

我们首先可以将各种各样的单个技术，定义为人类为了实现自己的目标而运用的手段。作为手段，技术可以包括工业生产程序、机械设备、医疗程序、算法法则以及商业流程等。除此之外，技术还可以包括组织机构、法律和制度，这些也都是人类实现自己目标时所用的手段。关于技术，非常重要的一点是，技术总是用零件、装配组件和半成品建构、组装、组合而成的。这些零件、装配组件等也是人类实现目标的手段，因此新技术是通过组合现有技术而形成的。[25]例如，激光打印机就是根据现有的激光技术、数字处理器及静电复印术创造出来的。处理器引导高度集中的激光束，在复印机硒鼓“印出”一个图像。我们可以把技术看成是这样一个系统：新的元素（即技术）不断地从现有元素中形成，而且这些新元素的存在又可能要求更新的元素出现。

其次，我们可以将经济定义为一系列的安排和活动，社会就是通过

这些安排和活动来满足自己的需要。这些安排当然就是指经济中的技术。从这种角度来看待经济也许并不常见，但是其实非常符合古典经济学家们的观点。因为古典经济学家认为，经济就是一个以自身的生产工具为起始点的过程。我们可以这样说，经济涌现于它自身的安排和自身的技术，经济就是它自身技术的一种表达。从这个视角而言，经济就是其自身的生产方式（即它自身的技术）的一个生态系统。在这个生态系统中，得到应用的各种技术必须是相互支持，在经济上要保持一致的。

在此基础上，我们还可以加入另一种观察结果。只有当存在对技术的“需求”时，技术才会出现。这种需求大多是来源于技术本身的需求。例如，汽车出现后，就需要或引发一系列更进一步的技术，包括石油勘探技术、石油钻探技术、炼油技术、大批量生产技术、汽油分销技术，以及汽车维修技术等。在任何一个时候，都存在着一个开放的机会之网，有利于进一步的新技术的开发和新安排的涌现。

基于以上内容，我们就可以搭建经济的基本结构了。要让这个基本结构运行起来，首先还要搞清楚技术集合是如何建构的。根据建构技术集合的各个步骤，我们可以给出如下经济形成的“算法”：

步骤1：新技术出现。新技术是在某些现有的特定技术的基础上创造出来的，并且它会作为一个新元素，进入当前的技术集合中。

步骤2：新技术变得活跃起来，并替代现有技术及现有技术中的某些部分。

步骤3：新元素为支持性技术和组织安排，创造出进一步的“需求”或提供进一步的机会利基（opportunity niches）。

步骤4：如果被替换的旧技术从技术集合中退出，那么它们的附属需求就会消失。它们提供的一系列机会利基，也会随着它们的退出而消失；反过来，那些用于填补这些机会之窗的元素，也会变得不再活跃。

步骤5：作为未来技术或未来元素的组件，新要素变得活跃起来。

步骤6：经济，即商品和服务的生产和消费模式，重新进行调整来适应上面这些步骤。成本和价格以及研发新技术的激励也会相应地有所改变。

例如，铁路机车是通过将现有的蒸汽发动机、锅炉、曲柄和铁轮组合而建构出来的。在1819年前后，铁路机车出现在了技术集合当中（步骤1）；它替代了原有的马车（步骤2）；创造了对铁路网络和铁路运输企业的需要（步骤3）；导致运河水运业和马车陆运业的萎缩（步骤4）；成为货物运输的一个关键组件（步骤5）；随着时间的流逝，最终使整个经济的价格和激励出现了变化（步骤6）。当然，以上这些事件或步骤可能是并行展开的，如新技术一出现，新机会就会随之出现。

如果你在脑海中依序“运行”这个算法，你就可以发现一些很有趣的事情。这个算法一旦运行起来，就可以启动一系列事件且永远不会停下来，因为这些事件中的每一个事件，都可能激发出更进一步的事件。例如，一项新技术可以通过步骤3和步骤5，导致更新的技术出现；通过步骤4，可以进一步替换旧技术；通过步骤6，则可以进行进一步的调整。同时，这些新技术反过来也会提供更多的机会、更新的技术，以及更进一步的替换。这个算法看上去也许很简单，但是一旦启动它，就会永不止息地引发各种各样的、具有一定模式的新行为。

至此，经济形成的基本机制也被解释清楚了。但是，除此之外还有第二个层面的机制，该机制能够进一步增加经济结构。新技术通常是以技术集群的形式进入经济的。在过去几十年间，一大批技术，如蒸汽发动机技术、电力技术、化学技术、数字技术等都已经进入了经济。所有这些技术，或者是以某项既定的关键核心技术为基础的，如蒸汽发动机技术是以蒸汽机为基础的；或者是以一系列相关现象为基础的，如相关的化学现象、电力现象、遗传现象等，这些现象得到了有效的利用，从而形成了相关技术。此外，这些技术是在以往的一两个核心技术的基础上逐渐形成的，即在核心技术的基础上，补充所需要的子技术。与其说这些“技术体”是在经济内被采用的，还不如说它们是与相关的行业“邂逅”的。它们与现有的商业流程相结合，创造出了新活动，激发出了新动力，形成了可用的新流程，并促成了小企业的急剧涌现。在这些小企业中，少数会继续成长为大企业。

经济，满足我们需求的一系列安排和活动，就是这样形成的。事实上，经济就是所有这些机制的运行结果。

截至目前，我介绍的这些机制，只是经济重塑自身的过程中最基本的“骨架”。每个机制都含有若干的子机制，它们在这里都被略去了。但是请读者注意一下，总的主题已经很明确地揭示出技术拥有一些简单属性，这些简单属性生成了一个由处于不断变化当中的元素（即技术）组

成的系统。在这个系统中，每个新的元素都是在以前元素的基础上形成的，每一种元素都会导致元素的替换。所有的元素都会带来一系列对于未来更新的元素不断变化的需求，而整个系统是在那些新近被发现的主导性现象，其属性和可能性的引导下结构化的。

这整个过程是一个自我创造的过程。新技术的形成来源于现有的技术，因此整个技术集合是自我生产的，或者自创生的（autopoietic）。经济也是自创生系统。经济的形成来源于技术，而且能够引发进一步的技术形成，进而导致经济自身的进一步形成。很显然，我们又一次深入复杂性领域的腹地了。

现在，我们可以了解到，经济的结构是怎样发生变化的。随着新的实体技术的进入，新的组织形式和新的制度也在新技术的“要求”下应运而生。而且，这些新的组织形式和新的制度，反过来也会对新技术产生更进一步的需求，即要求更进一步的方法、组织和制度的出现。于是，结构就涌现了出来。从更长远的时间尺度来看，大量的技术汇集成的技术集合，确定了经济的“主旋律”，即经济运行的主要方式。我们就是这样迎来了蒸汽时代、铁路时代和数字时代。

技术也提出了一系列特有的挑战或“主旋律性”的挑战，它们需要用新的方法去解决，经济也因此发生了结构性的变化。例如，蒸汽机和早期纺织机的出现，使得以工厂为基础的维多利亚式经济成为可能。不过，维多利亚式经济的过度发展，又引发了一系列新的安排，如保护儿童安全的法律、改善工作条件的法规，以及现代工会等。[26]随着经济的变化，经济组织和经济制度也随之发生了变化，而这些变化又会带来更进一步的安排、更进一步的新技术以及更进一步的变化。经济结构也因此发生变化。我们可以确定经济实现这种自我更新的机制，但是我们无法准确预测这些机制究竟以何种方式发挥作用。整个过程（如果你愿意的话，也可以称之为“计算”）远远不是确定性的，但是它确实可以说是一个非常完美的非均衡过程。

需要注意的是，我概述的这个理论实际上是一个算法而不是一组方程式。这个理论表现为一系列程序，其中一些程序由另一些程序触发。读者可能还会问，这怎么可能是一种理论呢？它当然是理论。事实上，它与生物学理论相类似。即便是在达尔文出版《物种起源》一书后的今天，也没有人能够成功地将新物种的产生、新物种生态系统的形成，以及进入某个特定物种主导时代的过程，归纳为一个方程组。原因就在于，进化过程是基于一些进化机制的，而这些进化机制是依序分步运行

并相互触发的，而且会不断地界定出新的类别（即新物种）。联立方程只有在类别给定时才大有用武之地，因为方程式能很好地描述出给定类别内的数值或数量变化，但是在新类别不断涌现的情况下，方程式通常就无能为力了。

我们必须承认，只有在深刻地理解进化的核心机制，并且提出一组与现实世界的现象相符合、内在一致的一般性命题时，我们这些“理解”才可以说真正构成了理论。[27]因此，生物学理论是理论，但不是以数学形式表达的理论；它是以过程为基础的，而不是以数量为基础的。总之，生物学理论是程序性的。同理，一个关于形成和变化而又周详的经济学理论也将是程序性的。[28]经济研究的要旨在于深入理解驱动经济形成的机制，而不在于能不能将这些机制归纳为方程式。我提出的这种程序性理论，并不是对标准理论的否定。但是这种理论确实提供了一种替代研究进路，它将重点集中在能够促成变化的动力本身，这就是技术。

然而，我们怎样才能更加深入地研究这些问题呢？这些基本过程都是有算法的，所以我们肯定可以针对它们的核心机制，构建一些基于计算机的模型。[29]本文介绍的这些研究只是一个开始。我们最终想要建构的总体理论，是一种关于创造性生成的理论：新元素是在现有元素的基础上形成的，新结构是在现有结构的基础上形成的，同时形成本身也源于之前的形成。这无疑是一个贯彻了复杂性思想的观点。

泡沫和崩溃是市场的基本趋势

行文至此，读者应该清楚，对于经济学问题，我们有了一种不同的思考框架。这种经济思想所强调的，并不是实物商品或现实的服务；相反，它强调的是变化和创造的过程。正如有的读者所想到的，这并不是一种全新的经济学思想，它与之前的经济学思想有一定联系。接下来，我将对这种联系加以说明。

经济学向来有两大问题。第一个问题是经济中的资源配置。所有市场中，商品和服务的数量与价格之间是如何被决定的？一般均衡理论、国际贸易理论和博弈论这些“伟大的理论”，都是研究这个问题的“典型代表”。第二个问题就是经济中的形成。经济最初是如何出现的？又是如何发展的？经济结构是如何随着时间的推移而变化的？关于创新、经济发展、结构变化、历史的作用、制度，以及治理的经济学思想，全都是研

究这个问题的结果。资源配置问题已经得到了深入的研究，而且这个问题已经高度数学化了，但是对于经济形成问题的了解却仍然非常少，而且这个问题也几乎完全没有被数学化。[30]

为什么会出现这种情况呢？直到1870年前后，所有伟大的经济学家都认为，这两个问题在经济学中同样重要。斯密、穆勒和马克思都努力尝试过，从理性科学的角度研究资源配置问题，试图在这方面有所建树，与此同时，他们对经济形成问题、治理问题和历史问题，也同样做出了重要贡献。此后，在维多利亚时代，出现了伟大的边际主义革命和一般均衡革命。在严格的理性假设和均衡假设下，它们使得资源配置问题转化为一个代数问题和微积分问题。但是，经济形成问题却无法进行这种转化。这是因为就经济形成问题的本质而言，它既不能被限制为静态的，也不能被限制为理性的。到了20世纪，数学化的理论被当成了唯一的“理论”，因此经济学的数学化大潮与经济形成问题擦肩而过。

马歇尔、凡勃伦、熊彼特、哈耶克和沙克尔等一系列经济学家，以及他们之后的许多制度主义者和历史学派经济学家，都对经济形成问题进行了研究。然而令人感到遗憾的是，在很大程度上，他们这些研究是针对特定历史时期、特殊化、基于案例和直觉主义的。总之一句话，它们是属于文字叙述型的，而不属于能够被一般化的理论推理。随着时间的流逝，这些被称为政治经济学的经济学理论，便被搁置到了一边。政治经济学理论虽然被公认为是符合现实的、有用的，但是却得不到大家的一贯尊重。

好在到了今天，经济学家们都非常清楚，关于资源配置的“数学分析”，不但远未包含所有经济问题，而且也不能很好地处理经济形成、探索、适应以及质变等问题。相比之下，复杂经济学则非常关注创造问题和结构形成问题，它研究的就是创造和结构形成得以实现的机制。复杂经济学是与政治经济学一脉相承的，并且是对政治经济学伟大传统的复兴。特别让我高兴的是，这两种理论有很多值得相互借鉴的地方。利用复杂经济学，我们能够从理论上系统地研究世界的形成；利用政治经济学，我们能够通过直觉方法和经验方法来研究世界的形成。复杂经济学将为政治经济学提供坚实的理论基础。复杂经济学不可以、也不应该取代基于案例的历史分析；恰恰相反，它将深化和发展这个值得尊重的经济学思维方法。同样，政治经济学也会深化和发展复杂经济学。

政治经济学的一个主要优势，就是它的历史感和它的时间感。时间可以导致真实的、不可逆转的差异，进而不断创造出新的结构。相比之

下，新古典主义经济学在处理时间问题时，则要逊色很多。在均衡状态下，一个结果会一直持续下去，因此时间在很大程度上没有多少意义；相反，在动态模型中，时间却变成了一个重要的参数，这个参数可以来回转换、影响当前的结果。这一点令很多经济思想家觉得“不舒服”。1973年，琼·罗宾逊（Joan Robinson）提出了一个著名的主张：“一旦我们承认经济存在于时间之流中，承认历史是朝着一个方向发展的，即从不可逆转的过去向不可预知的未来发展，那么均衡概念也就站不住脚了。到那时，我们就要重新思考整个传统经济学了。”

在反思时间问题这个方面，复杂经济学与政治经济学是一致的。在“计算”中，也就是说在经济中，大概率事件和小概率事件，往往会在某些特定的、不可重复的时刻决定吸引子的形成，决定时间结构的形成和消失，决定技术的创生，并决定源于这些事物的经济结构和制度。而技术和技术结构，反过来也建立在这些事物之上，甚至未来的经济形态，即通往未来的道路，也是建立在这些事物之上的。无论在什么层面，无论在什么时候，经济都具有路径依赖性。因此，历史再次变得至关重要，而时间也得以再次出现在经济中。

人们可能提出的一个很自然的问题是，复杂经济学这种经济学新框架是否有它的政策含义。当然有啦！复杂经济学告诉我们，市场自身就有出现泡沫和崩溃的趋势。市场会诱发多重局部吸引子状态，市场能通过金融网络传播各种事件，市场能创造出一系列技术解决方案，同时又带来一系列挑战。有了这样的认识，就不难了解哪些政策大有用武之地了，如监管过度行为的政策、以“轻推”方式促成有利结果的政策、创造有利于创新条件的政策等。科兰德和库珀尔（Kupers），对这些政策有一个很好的概括：创造适宜的元条件（meta condition）。

上面这种说法当然没有错。但是我认为，我们还可以提出一个更加强有力的观点。新古典经济学之所以在现实世界中遭到失败，很大程度上是因为它把经济看成是均衡的。如果我们回头看看过去25年间发生的历次经济危机，就会发现：所有这些危机在很大程度上，都是由于少数处于有利地位的“大玩家”利用经济系统中的漏洞而造成的，或者说是由失控的市场所导致的。例如，1990年，俄罗斯在放开市场之后的经济崩溃；2000年，美国政府放松监管之后，加利福尼亚能源市场危机爆发；2008年，冰岛银行倒闭、持续的欧元危机，以及华尔街危机等。受均衡思维所限，人们无法提前预见到这种情况。原因很简单：由定义可知，在均衡状态下，没有人有动力偏离当前的行为，所以“钻空子”行为不会

发生。受限于均衡思维，人们也很难看到极端的市场行为，因为他们认为对均衡的偏离，很快就会被反向抗衡力量纠正过来。从均衡理论的基本假设就可以看出，研究经济的某些组成部分如何被钻空子、研究经济的系统性崩盘，本来就不是均衡经济学的主要目的。

相比之下，复杂经济学却可以告诉我们，经济系统永远对各种各样的反应开放，它的每个组成部分永远对各种各样的新行为开放，包括通过钻空子来谋利、结构上的突然变化等。复杂经济学建议出台精心设计的防控措施，就像政府为地震多发地区制定合理的建筑安全标准一样。而且，同样重要的是，复杂经济学会让我们转向现实主义立场。经济并不是由一系列没有任何激励去改变、自然就能共同创造最优结果的行为所构成的。恰恰相反，经济就是一张激励之网，它总能激发出新的行为，诱发新的策略，并让它们共同形成“合理”的结果，从而驱动系统不断变化。

一门基于“动词”的科学

> 均衡系统（**Equilibrium System**） 一些经济学家认为，均衡就是经济的自然状态。这时，经济体系中各种相互关联和相互对立的因素处于相对平衡和相对稳定的状态。

复杂经济学不是添加到标准经济学上的附加组件，它也不意味着简单地将基于行为主体的行为加入到标准模型中去。相反，复杂经济学以一种不同的方式来思考经济。它从来不将经济视为一个均衡系统，而是把它看作一个动态系统，一个不断进行自我“计算”的系统，一个不断自我创建、自我更新的系统。均衡经济学强调秩序、确定性、演绎推理和静态均衡，而复杂经济学强调偶然性、不确定性、意义构建（sense making）和一切变化皆有可能。或者换一种说法，迄今为止，经济学一直都是一门基于“名词”的科学，而不是一门基于“动词”的科学。经济函数随时间的变化，通常都被定义为固定的名词，即实体层面的变化，如就业、生产、消费、价格等。但是现在，这些变化已经从名词实体层

面，转换到了动词行动层面，如预测、反应、创新、替代等。行动能够引发进一步的行动。

这一转变深刻揭示了中间层经济，即中观经济在经济中的重要地位，同时也重新定义了经济学中的问题的解。从此，问题的解不再是一组数学条件，而是一种模式、一系列暂时现象、一系列能够引发进一步变化的变化、一系列能够创造新实体的现有实体。理论研究的目的，也不再是发现那些“不朽”的一般性定理，而是更深刻地理解创造出这些模式、并使这些变化得以传播的机制。

从更广阔的角度来看，经济学的这个转变，也是科学本身大转变的一个非常重要的组成部分。与以往相比，现在所有科学都在变得更加程序化、算法化、“图灵化”，所有科学都更少依赖方程式、连续性、“牛顿式”了。之所以会出现这种趋势，主要由于以下两个原因：一是生物学作为一门严格科学的兴起，二是计算和计算机科学的崛起。即便是数学本身，也在向这个方向转变。例如，格里高利·蔡廷（Gregory Chaitin）曾经指出，数学正在从连续的公式、微分方程、静态的结果转向离散的公式、组合推理及算法思维。他说：“计算机不仅是一种极其有用的技术，还是一种具有革命性意义的新数学，它带来了深刻的哲学后果，它揭示了一个新世界的面纱。”科学和数学中的确定性正在减少，它们正在走向开放、拥抱程序性思维。在这个方面，经济当然也不能例外。

复杂经济学也不是新古典经济学的一个特例。恰恰相反，均衡经济学是非均衡经济学，也就是复杂经济学的一个特例。我们可以说，复杂经济学是用更具一般性的方法来研究经济学的。当然，均衡仍然是一个非常有用的一阶近似。在解决经济中可以明确界定的、可以理性化和静态的问题时，它是有用的。但是，不能再将均衡经济学称为经济学的中心。稳步走向经济学中心的是复杂经济学，[31]它能够更一般地处理相互作用，能够承认非均衡现象，能够处理创新、形成和变化等问题。

当然，复杂经济学仍处于发展的早期阶段，很多经济学家都在努力扩大复杂经济学的研究“领地”。复杂经济学告诉我们，经济永远都在发明自身，永远在利用机会创造可能性，永远在应对各种变化。经济不是死的、静止的、永恒的和完美的，恰恰相反，经济是活的、永远处于变化之中的、有机的和充满活力的。

COMPLEXITY AND THE ECONOMY

02

“爱尔法鲁酒吧”问题

归纳推理和有限理性

“爱尔法鲁酒吧”问题是个很有意思的决策问题。“很多人会去”的预期会导致几乎没人去，而“不会有人去”的预期则会导致许多人去。为什么理性预期会导致自我否定？现代心理学家认为，在复杂的或不确定的情况下，演绎推理会显得“力不从心”，这时人们主要依靠归纳推理进行决策。

“爱尔法鲁酒吧”问题（**El Farol Problem**） 布莱恩·阿瑟1994年提出了一个著名问题：关于去酒吧还是不去酒吧，人们究竟是如何进行决策的？

1993年，当时我在圣塔菲研究所，正尝试用计算机模拟的方法来对以归纳方式进行经济决策的问题进行建模，不经意间发现了一个悖论。在圣塔菲的峡谷路，有一家名为“爱尔法鲁”的酒吧。每个星期四晚上，酒吧都有表演。人们如果预期那里人不多，他们就会去；如果预期那里人很拥挤，他们就不去。我马上意识到，这是一个很有意思的决策问题。“很多人会去酒吧”这种预期会导致几乎没人去，而“不会有什么人去酒吧”这种预期则会导致许多人去。这就是说，预期所导致的结果，恰恰否定了预期。或者更准确地说，这意味着理性预期，即一般来说正确或有效的预测，会导致自我否定。因此，这里出现了逻辑上的自我矛盾，而且与说谎者的悖论相比，逻辑结构不一样。

我据此写成了一篇文章，得到了许多物理学家的激赏。在复杂性研究的圈子中，“爱尔法鲁”变成了众所周知的名字。“爱尔法鲁”问题后来被张翼成和达米安·夏利一般化了，并进入了博弈论教科书。许多学者都对这个问题进行了研究，撰写了很多文章，为最初的问题和后来提出的少数者博弈模型及其变体给出了很多个“解”。

这篇文章最初发表于1994年的《美国经济评论会议文章》上。

* * *

经济学中假定的理性类型，是个完美的、合乎逻辑的、演绎的理性，这在对理论问题求解时非常有用。但是，对人类行为的这种理性假设是有很大问题的，它带来的问题远远多于它通常能够解决的。如果我们把经济主体所面临的决策问题想象成一个大海，即简单的决策问题构成表层和浅层，复杂的决策问题构成深层和底层，越靠近海平面的决策问题越容易，那么演绎理性最多只能解决位于海平面及以下1～2米的那些问题。例如，“井”字游戏问题很简单，它位于上述“决策问题之海”的表层，很容易就可以为它找到一个合乎完美理性的极小极大“解”。但是，稍稍复杂一些的国际跳棋，它位于“决策问题之海”略深一点的地方，就找不到理性的“解”了；至于国际象棋和围棋，仍然位于“决策问题之海”的中等深度的地方，当然就更加找不到了。

完美理性或演绎理性，在面对复杂情况时必定会“力不从心”，这有两个原因。第一个原因是显而易见的：当复杂性超过了一定程度时，人类的逻辑思维能力就无法应付了，这就是说，人类的理性是有限的。第二个原因是：在多个行为主体相互作用的复杂环境下，任何一个行为主体都不能假设与自己互动的其他行为主体的行为是完全理性的，因此每

个行为主体都不得不猜测其他行为主体将如何行动。这就是说，他们必须依据主观信念以及关于主观信念的主观信念来做出决策。因此也就不存在客观的、明确的、共同的假设了。反过来说，这也就意味着，理性的、演绎性的推理，即在明确的前提假设下，通过完美的逻辑推理得出的结论不再适用了。总之，问题本身将变得晦暗不明。

有限理性（**Bounded Rationality**）　介于完全理性和非完全理性之间的、在一定限制下的理性。这个概念最初是阿罗提出的，后来，赫伯特·西蒙提出了有限理性模型，有限理性决策理论应运而生。

当然，经济学家对这些情况其实是非常清楚的。问题不在于完美理性有没有效，而在于完美理性在什么地方有效。如果承认有限理性，那么在经济学中，如何对有限理性建模？关于有限理性的文献虽然不太多，但一直在不断增加，它们蕴含着很多有价值的思想，但是这些思想在很大程度上是零碎的，彼此之间没有很好地衔接起来。相比之下，在行为科学中却不是这样。现代心理学家基本上已经有了这样一个共识：在复杂的或不确定的情况下，人们会使用一些有自身特点的、可预测的推理方法，而且这些推理方法都不是演绎性的，而是归纳性的。

归纳思维

在复杂或不确定的情况下，人们是怎样进行推理的？现代心理学告诉我们，作为人类，我们只拥有适度的演绎推理逻辑能力，而且我们只能适度地运用这种能力。但是，我们却特别擅长观察、识别和匹配模式，也就是那些能够带来明显进化利益的行为。因此，我们在面对复杂问题时，我们会先搜寻模式，并利用找到的模式来简化问题，然后构造临时的内部模型或假说，或者说“图式”（schemata）。[1]接着，我们会根据当前的假说来进行局部演绎推理，并采取行动。当我们接收到来自环境的反馈后，我们对当前加强假说的信念可能会强化，也可能会弱

化；那些没有用的假说将会被丢弃，并在需要时用新的假说来替代。换句话说，当我们不能完全依靠演绎推理，或无法对问题进行完备的界定时，我们就使用简单的模型，来填补我们理解中的空白。这种行为是归纳性的。

演绎推理（**Deductive Reasoning**） 就是从一般性的前提出发，通过推导，即“演绎”，得出具体陈述或个别结论的过程。

在下国际象棋的棋手身上，我们可以观察到这种“归纳性的行为”。一般来说，在下棋时，棋手会研究整盘棋的“局面”或“大势”，并回忆自己的对手在过去比赛中的下法，以便识别不同的模式。然后，他们会利用识别出来的模式，来形成关于对方意图策略的假说或内部模型。而且，在同一时刻，他们的头脑中可能会“准备”好几个内部模型，如“他正在使用‘卡罗–坎恩’（Karo-Kann）防御”“这看起来有点像博特温尼克（Botvinnik）和维德马（Vidmar）在1936年那个对局”“他可能想走出中盘兵布局”等。棋手们在这些假说的基础上，进行局部演绎推理，分析不同的下法可能会带来的影响。随着棋局的展开，他们会坚持那些被证明有效的假说或心理模型，舍弃那些被证明无效的假说或心理模型，并且形成新的假说或心理模型。换句话说，他们在下棋时要完成一系列“工作”：识别模式、构造假说、根据当前持有的假设进行推理，并随时根据需要替换假说。

这种类型的行为在经济学中可能不为熟知，但是我们要认识它的优势其实并不困难。采取这种方法，我们就能够解决复杂问题。我们可以构建可行的且更加简单的模型，那是我们可以处理的。它还使我们能够处理不明确的问题，在界定不清楚的地方，我们的工作模型能够自动填补空白。它不是“理性”的对立面，当然更不是科学的对立面。事实恰恰相反，科学本身的工作方式和取得进步的方式就是如此。

对归纳推理建模

如果人类真的是这样推理的，那么我们应该怎样做，才能将这种推理模型化呢？如果决策问题是随着时间的推移来完成的，那么一个很自

然的思路就是：设定一个由行为主体组成的集合，行为主体可以是异质的，同时也假定他们可以形成各自的心理模型、假说或主观信念。这种主观信念可以用不同形式来表示。它们可以用简单的、用来描述或预测一些变量和行动的数学方程式来表示，或者可以用经济学中常见的那种形式相当复杂的预期模型来表示，又或者可以用统计假说或“条件-预测”规则来表示，例如，“如果观察到情况Q，那么就预测结果或动作D”。所有这些通常都是主观的，也就是说，它们随行为主体不同而不同。一个行为主体在某个时刻，可以只持有一个主观信念，也可以同时持有多个主观信念。

每个行为主体，都需要将自己的各个信念模型的“历史绩效”记录下来。当需要做出选择时，他就根据当前他认为最可信的，或者说可能是最有利可图的那个信念模型来采取行动。至于其他信念模型，他先记在心里。或者，他也可以根据若干个信念模型组合来采取行动。然而，人类的一般倾向是，记住许多信念模型或假说，可以根据最合乎情理的那一个采取行动。一旦采取了行动，整个图景就会得到更新，同时行为主体也会更新所有假说的“历史记录”。

这是一个“会学习”的系统。通过“学习”，行为主体知道他们的假说当中哪些是有效的，他们会不时舍弃“表现不佳”的假说，并生成新的“想法”以便取而代之。行为主体坚持当前最可信的假说或信念模型，但是一旦它不再有效，他们就会放弃它，转而采用一个更好的。当然，这会导致一种内置的滞后。一个信念模型之所以被坚持，不是因为它是“正确的”，行为主体没有办法知道它是否正确，而是因为它在过去是有效的，因此在认定它应该被舍弃之前，必须先积累一些关于它的“失败”记录。一般而言，每个行为主体都可以说，有一个据此采取行动的“缓慢周转的假说库存”。或者，我也可以说这是一个临时得到了实现的预期系统，当然只是暂时实现，而不是完美地实现。当这些信念、模型或假说不能再实现时，它们就要让位给不同的信念、模型或假说。

有的读者可能觉得自己对这个系统不熟悉，那样的话，他或她或许可以回想一下经济学中标准的学习模型，然后把这个系统视为它的一般化。在标准的学习模型中，所有行为主体通常共享一个具有未知参数的预期模型，并根据当前最合理的参数值来决定自己的行动。与标准的学习模型相比，我们的行为主体是异质的，并且每个行为主体都运用若干个主观模型，而不是共享一个统一的模型。显然，这是一个更加丰富多彩的世界。人们可能会问，在特定的情况下，它是否会收敛到某种标准

的信念均衡，或者会不会永远保持开放，总会产生新的假说和新的想法。

这也是一个进化的世界，或者更准确地说，这是一个共同进化的世界。这就像生态系统中的物种一样，为了生存和繁殖，必须通过竞争、通过适应自己和其他物种共同创造的环境来证明自己。在我们这个世界里，各种假说要想变得准确、要想成为行动的依据，就必须通过竞争、通过适应自己和其他行为主体的假说共同创造的环境来证明自己。因此，在任何一个阶段，作为行动依据的一系列想法或假说将共同进化。[2]

然而，一个关键问题仍然悬而未决。假说或心理模型到底来自哪里？或者说，它们是怎样生成的？从行为的角度看，这是心理学中的一个深层问题，与认知、对象表征和模式识别有关。我在这里不可能深入探究这个心理学问题。不过，在构建模型时，还是有不少既简单又实用的方法。有时，我们可以赋予行为主体某些焦点模型。焦点模型指的是那些明显的、简单的和容易处理的模式或假说。例如，我们可以生成一个“焦点模型库”，然后分配给行为主体。而在其他时候，在给定一个合适的模型空间的情况下，我们可以用遗传算法或其他类似的“智能搜索设备”，生成“更有智能”的模型。或者，我们还可以允许行为主体相互“借用”心理模型。在心理学中，这被称为模型“迁移”。无论选择哪种建模方法，重要的是一定要明确，上述框架必须独立于所使用的具体假说或信念，就像消费者理论框架独立于所选择的特定产品一样。当然，要在特定问题中应用某个框架，则必须采用某个生成信念的系统。

60人！归纳推理揭开的谜底

> 归纳推理（**Inductive Reasoning**）　在某些决策环境中，我们先形成各种各样的工作假设，并根据其中最可信的那个工作假设采取行动，如果不再有效，那么就用新的工作假设取而代之。

接下来，我将构造一个问题，以它为例来说明何为归纳推理，并阐述如何对归纳推理建模。假设现有N个人，每个人独立决定要不要在某个晚上去酒吧消遣。为了更具体起见，我们不妨假设N为100。酒吧的空间是有限的，如果酒吧里不太拥挤，那么来酒吧的人就可以度过一个愉快的夜晚。具体地说，如果这100个可能去酒吧的人当中，只有不到60个人真的去了，那么去的那些人就可以度过一个愉快的夜晚。没有人能够提前知道当晚会来酒吧的确切人数。因此，如果一个人或行为主体预期，当晚去酒吧的人小于60个，那么他就会去（他认为值得一去）；如果他预期去酒吧的人会超过60人，那么他就宁愿待在家里。这些人在做出决策时，不会受自己以前有没有去过酒吧的经验的影响，同时不同行为主体之间也不存在共谋或事先的沟通。唯一可用的信息，是过去几个星期以来出现在酒吧中的人数。这个问题的灵感，来自圣塔菲研究所旁边一家名为“爱尔法鲁”的酒吧。这家酒吧每周四晚上都有爱尔兰音乐会。不过，很多类似的场所中都会出现同样的问题，比如读者可以想象一下你去吃午餐的餐厅，你也许希望它安静点，但是它可能很拥挤。事实上，任何“公地问题”或“协调问题”，只要涉及数量上的限制都是一样的。在这个问题中，我们感兴趣的是，每个星期来酒吧的人数变化的动力学机制。

不难看出，这个问题有两个有趣的特点。首先，如果有一个“显而易见”的模型，所有行为主体都能够根据这个模型，预测来到酒吧的人数并在此基础上决定去不去酒吧，那么通过演绎推理就能够求解了。但是，这里的问题显然不属于这种情况。给定最近来酒吧的人数，可以设想一大批看上去同样合理、同样有根据的模型。因此，也就无法得知其他行为主体可能选择哪个模型，这样某个“有参照意义”的行为主体，也就不能以某种确定的方式给自己选定一个模型。这里不存在演绎理性解，即没有“正确的”预期模型。从行为主体自身的角度来看，这里的问题是不明确的，因此他们被推进了一个归纳的世界。其次，在这里令人烦恼的是，任何一个共同的预期都会被打破。如果所有人都预测很少有人会去，那么所有人都会去，而这个结果将证明这种信念是无效的。同样地，如果所有人都预测大多数人会去，那么将没有人会去，这种信念同样被证伪。[3]由此，行为主体们的预期将被迫变得有所不同。

行文至此，我想请读者暂且先停下来思考一下：随着时间的推移，来到酒吧的人的数量（为了行文方便，以下简称为“到场人数”）会如何动态地变化。它会不会收敛？如果会收敛，那么为什么收敛？或者，它会不会陷入混沌呢？我们又该怎样进行预测？

动态模型

为了回答上述问题，我将根据上面描述的框架构建一个模型。假设这100个行为主体，每个人都可以形成若干个“预测器”或假说，即有这样的函数，将过去d周的到场人数映射为下周的到场人数。例如，最近各周的到场人数可能构成了如下序列：

……44，78，56，15，23，67，84，34，45，76，40，56，22，35。

据此，行为主体构建的假说或“预测器”可能是这样的，预测下周的到场人数将为：

- 与上周的到场人数一样［35］，
- 与上周的到场人数以50为中心构成镜像［65］，
- 过去4周到场人数的（四舍五入）平均值［49］，
- 过去8周到场人数的趋势，上下界为0和100［29］，
- 与两周之前的到场人数相同（以两周为周期的循环）［22］，
- 与5周之前的到场人数相同（以5周为周期的循环）［77］，

……

再假设，每个行为主体拥有一个由k个这样的“焦点预测器”组成的集合，而且可以将这个集合的“行为”记录下来。这样一来，他就可以用这个集合中当前最准确的那个“预测器”，来决定是去酒吧还是留在家中。我把这个“预测器”称为他的“活跃预测器”。一旦做出了决定，每个行为主体就会知道新的到场人数，同时更新他那些“预测器”的准确程度。

需要注意的是，在这个“酒吧问题”中，尽管决定到场人数的是行为主体们在采取行动时所根据的当前最可信的假说的集合，即“活跃假说”的集合，但是“活跃假说”的集合又是由历史到场人数所决定的。用约翰·霍兰德给出的术语来说，我们可以认为这些“活跃假说”形成了一个“生态”。我们感兴趣的是，这个“生态”如何随着时间的推移而进化。

计算机实验

对于大多数假说的集合来说，要得出解析解都很困难，因此我在下

面将通过计算机实验来求解。在计算机实验中，为了生成假说，我先创建了一个预测器的“字母汤”，方法是将十几个焦点预测器连续复制很多次。然后，我随机地将k个（比如说，6个、12个或23个）预测器，赋予这100个行为主体中的每一个。这样一来，每个行为主体就拥有了k个预测器（或假说，或他可以利用的“想法”）。我们不必担心没有用的那些预测器会使得行为主体无从选择。如果预测器是“无效”的，那么它们就不会被使用；如果它们是“有效”的，那么它们就会被放到最前面来，成为“活跃的”预测器。在这个问题中，当给定了初始条件和每个行为主体可以运用的预测器集合后，所有预测器在未来的准确度都是预先确定下来的。因此，这里的动力学是确定性的。

实验结果非常有意思（如图2-1所示）。在出现了“循环探测”预测器的情况下，这种循环很快就因为被“套利”而消失了，因此不会存在持续的循环。如果有几个人预测，因为三个星期之前有许多人去，所以下个星期也有很多人去，那么他们将会留在家里。更有趣的是，平均到场人数总是会收敛到60人。这些预测器通过“自我组织”，形成了一个均衡模式或“生态”。在这个“生态”中，平均来说，所有活跃预测器，即那些最准确的、行为主体据此采取行动的预测器当中，有40%预测到场人数高于60人，60%预测到场人数低于60人。

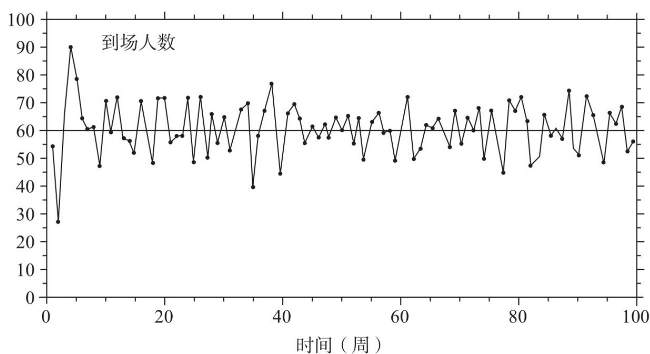


图**2-1**　爱尔法鲁酒吧前**100**周的到场人数

从根本上看，这种涌现出来的“生态”可以说是有机的。这是因为，活跃预测器总体上可以按60：40分成两类，每一类的成员都在不断地变化。这就像一个森林，它的轮廓不会改变，但是组成它的每一棵树却一直在改变。这些结果在所有计算机实验中都出现了，而且当预测器类型以及分配给每个行为主体的预测器数量改变时，结果仍然非常稳定。

那么，这些预测器是如何实现自组织，导致了到场人数平均为60

人，预测结果则按60：40的比率来分类的呢？一个可能的原因是，在这个“酒吧问题”中，60也许是一个自然的“吸引子”。如果我们把这个问题视为一个纯粹的预测博弈，那么不难看出，以40%的概率预测到场人数多于60人，以60%的概率预测到场人数小于40人，这个混合策略正是一个纳什均衡。然而，这仍然不能解释行为主体是如何接近这样的结果的，因为他们的推理是现实的、主观的。为了更好地理解这一点，不妨先假设70%的预测器在相当长的时期内，预测到场人数都在60人以上。这样一来，实际平均到场人数却只有30人，从而“证实”了到场人数接近30人的预测，并“证实”了到场人数高于60人的预测是不成立的。这将有助于恢复预测器之间的“生态平衡”，最后会调整为40%～60%的组合，而这个组合是可以“自我维持”的。但是，要从数学上精确地证明这一点，似乎并非易事。重要的是，我们一定要认识到，在设置预测器时，根本不需要考虑40～60人的“平衡预测”。尽管有的可能倾向于预测到场人数多，有的可能倾向于预测到场人数少，但是在总体行为中，上述“平衡预测”肯定会脱颖而出。当然，如果硬性规定所有预测器都只能预测到场人数低于60人，那么这个实验就会失败。结果是，所有100个行为主体总是都会到场。预测器必须能够在一定程度上“覆盖”整个预测空间。读者不妨思考一下，如果规定所有行为主体共享同一组预测器，结果会发生什么。

可能有人会反对说，在这些计算机实验中，我所使用的固定的、笨拙的预测模型给行为主体的预测造成了妨碍。如果他们能形成更开放的、更智能化的预测，那么不同的行为模型就可能涌现出来。当然，这可以算是一个猜想，有兴趣的人可以利用某种更加复杂的方法，例如遗传编程来检验这种猜想。遗传编程方法能够不断形成新的假说、新的预测器，它们能够利用“智能化”进行调整，并且随着时间的推移会变得更加复杂。我相信，我在上面给出的结果不会有任何质的变化，但如果在性质上真的被改变了，我会觉得非常惊讶。

对于这里介绍的“酒吧问题”，可以用很多方式加以推广。我也建议读者自己动手尝试进行计算机实验。

归纳推理的多彩世界

上面描述的归纳推理系统包括了多个“元素”，即信念模型或假说，这些“元素”要适应于它们共同创建的总体环境。因此，这个系统可以说

是一个自适应复杂系统。在经过一段初始的学习之后，行为主体所使用的假说或心理模型相互适应下来。因此，我们可以将一组相互一致的心理模型视为一组能够在一定条件下很好地协同发挥作用的假说。这体现了很高程度的相互适应性。有时，系统中存在着一个独一无二的这种集合，它对应于一个标准的理性预期平衡，所有信念都被它所吸引着。不过，在更通常的情况下，存在着多个这样的集合，这种可能性要高得多。在这种情况下，我们就有理由认为，经济中的归纳推理系统，无论是在股市投机、谈判、扑克游戏、寡头定价中，还是产品的市场定价中，都会陷入循环或暂时锁定在某些心理模式中。这些心理模式可能是不会重复出现和路径依赖的，而且可能会越来越复杂。总之，可能性非常多。

长期以来，对于复杂的、可能无法界定清楚的决策环境，经济学家依然假设完美的演绎理性，但是他们又会为此感到不安。人类应用完全理性的范围，最多只能称得上还可以。这种说法也许很令人惊讶。到目前为止，我们还不清楚，应该如何处理不完全理性或有限理性。然而，上文对“酒吧问题”的分析表明，人类在这类决策环境中，运用的是归纳推理：我们形成各种各样的工作假设，并根据其中最可信的那个工作假设采取行动，如果不再有效，那么就用新的工作假设取而代之。对于这种推理，可以采用多种方式进行建模。通常，这种建模会引导我们进入一个丰富多彩的心理世界。在这个世界中，每一个行为主体的“想法”或心理模型，都要与其他行为主体的“想法”或心理模型竞争，以求得生存。

这就是一个进化的、复杂的世界。

COMPLEXITY AND THE ECONOMY

03

圣塔菲人工股票市场

内生预期的资产定价

资产定价理论认为：行为主体是完全异质的，他们的预期需要不断适应市场，而市场本身则是他们的预期共同创造的。实验表明，在归纳预期的股票市场上，市场行为显著地偏离了理性预期均衡，真实股票市场上的泡沫和崩盘等特征也都一一出现。

圣塔菲人工股票市场（**the Santa Fe Artificial Stock Market**）　布莱恩·阿瑟通过计算机技术建立的一个人工股票市场，用来模拟现实中的股票市场。有意

思的是真实股票市场上的泡沫和崩盘等特征在这个人工股票市场上都一一出现了。

这篇文章源于20世纪80年代末，我们在圣塔菲研究所完成的一系列计算机仿真实验，当时我们正在研究基于行为主体的建模方法。在那个时代，标准的新古典股市模型的假设是这样的：所有投资者都是同质的，他们对股市涨跌的预测（预期）相同，而且他们的预测平均来说是正确的。虽然这种理论看上去很优雅、很精美，但是它的假设是不现实的。此外，它不仅排除了股市出现泡沫和崩盘的可能性，而且也排除了技术交易（即利用过去的价格模式来预测股票趋势）、价格和波动程度的相关性，以及大成交量交易等现实股票市场中会出现的现象。我们在计算机上构建了一个不同于标准的新古典股市模型的模型。在我们的模型中，“投资者”是一些异质的“人工模拟行为主体”，即小型计算机程序，他们必须提出自己的预测，而且必须搞清楚哪些预测是对的，哪些预测是不对的，当他们获得了市场经验后，就能做到这一点。

我们的人工股票市场实验的结果表明，在设定了一组符合现实的参数后，它就成了一个“复杂的股票市场体制”，泡沫和崩盘、技术交易、价格和波动程度的相关性及大成交量交易，所有这些现象都出现了。而在一组更狭隘且不现实的参数条件下，它也可以展示一个“新古典意义上的股票市场体制”，此时，标准理论仍然成立。这篇文章是我和约翰·霍兰德、莱克·勒巴伦、理查德·帕尔默、保罗·泰勒等人合力撰写的，最初发表在1997年的《经济学评论》杂志（*Economic Notes*）上，后来又被收入了由我、戴维·莱恩和史蒂文·杜尔劳夫（Steven Durlauf）主编的《经济可看作是进化的复杂系统》第二卷（*The Economy as an Evolving Complex System*II）中。[1]我们的模型被称为“圣塔菲人工股票市场”，它有很多变体，现在已经被广泛应用于经济学研究。

* * *

学术界的理论家看待金融市场的方式，与市场上的交易者通常存在着十分惊人的差异。标准的有效市场金融理论假定：投资者是完全同质的，他们对资产的未来价格有共同的理性预期，并且能非常理性地立刻将所有市场信息反映到这个价格中去。[2]由此得到的结论是，投机者不

可能获得投机利润。除非运气很好，否则无法通过技术交易（即利用过去的价格模式来预测未来价格趋势）来获利。临时性价格的过度反应，如泡沫和崩溃，反映出的是资产估值的合理变化，而不是投资者情绪的突然变化。其他一些结论还包括：交易量会很低或者为0，交易量和价格波动之间不存在任何序列相关性等。总之，根据这个标准理论，市场是理性的、机械的和高效的。

市场是理性的吗

与理论家不同，交易者则经常看到：市场提供了很多投机机会。在许多交易者看来，技术交易是有利可图的，[3]“市场心理”这种东西确实存在，与市场新闻没有关系的“羊群效应”可能会导致泡沫和崩溃。一些交易者和财经记者甚至认为，市场有自己的情绪和个性，他们有时会将市场描述为“紧张的”“迟钝的”或“不稳定的”。据这种观点看来，市场是“有心理的”、有机的和不完全有效的。而理论家们则认为，持这种信念的交易者是不理性的，甚至是迷信的。当然，这些交易者使理论家们假设的理性行为主体很“尴尬”，因为从交易者的角度来看，标准的经济学理论是不现实的，他们的看法与这种理论完全冲突。[4]

虽然没有什么理论家会说市场是“有个性的”、会经历情绪的起伏，但是标准的经济学观点近年来确实已经开始有所改变了。1987年的股市崩溃，沉重打击了经济学家坚持的信念。突然的价格变化，反映了对市场新闻的“理性调整”。研究未能发现，股市崩溃与当时发布的市场信息之间存在着显著的相关性，却发现股票市场中的交易量和价格波动都很大，而不像标准理论所预测的那样很小或者为0，并且具有显著的自相关性。股票的收益也包含着虽说不太大但很显著的序列相关性。而且，利用某些技术交易规则，能够带来统计上显著的、不仅仅是“适度”的长期利润。

长期以来，经济学家都知道，如果投资者将完全理性规则应用到证券市场中去，那么他们就会缺乏参与交易和收集信息的激励。如今，我们已经积累了足够的统计证据，足以证明有效市场理论是有疑问的，交易者的观点是不容忽视的。因此，现代的金融学文献一直都在寻找一种可以解释真实市场的理论，以此替代有效市场理论。

一个很有前途的替代理论是噪声交易者理论。它认为，当市场上存在“噪声交易者”，即其预期不同于理性预期交易者的那些投资者时，技

术交易策略，如趋势交易法就可能是合理的。例如，噪声交易者认为股价将持续上涨，那么理性交易者就可以按照上升趋势买入来获利，从而进一步强化了这种趋势。根据这种理论，正反馈型的交易策略，以及其他技术交易策略可以被视为是理性的，只要市场中存在着非理性交易者，就可以启用这些策略。

在一定程度上，这种噪声交易者“行为”理论是在向交易者的观点靠拢。但是，它仍然建立在两个不太现实的假设的基础上。第一，市场上存在着一些不聪明的噪声交易者，他们不会学习，因此总是做出错误的预测。第二，市场上存在着理性交易者，他们通过理论家没有明确指出的途径，拥有噪声交易者的预期和自己的预期的全部知识。在现实的证券市场上，这两个假设显然都不可能成真。为了说明这一点，假设某一刻市场上有一些只拥有最低智能的噪声交易者。随着时间的推移，肯定有部分噪声交易者会发现自己的错误，并开始形成更加聪明一些的，或者至少是不同的预期。这样一来，市场就会发生变化，而这就意味着拥有“完美智能”的交易者必须重新调整自己的预期。但是，没有任何理由认为这些“聪明人”知道噪声交易者偏离后的新预期。他们必须通过某种手段，比如说猜测或观察市场来形成新的预期。随着这些理性市场参与者的改变，市场将再次改变。因此，噪声交易者的预期可能再次进一步偏离原预期，从而迫使理性交易者进一步调整。尽管理论上假定噪声交易者的市场是固定不变的，但实际上这个市场将开始“散开来”，因此完全理性的交易者在每一轮的变化中都将通过观察市场来猜测改变后的预期。

噪声交易者理论虽然能解释一些现象，但仍然远远称不上一个稳健的理论。不过，幸运的是，通过反思这个理论，我们可以得到一系列有趣的想法。假设我们也假设了“理性”，但是假设行为主体是不同的，他们并没有发现自己在一个有理性预期的市场中，或者说他们并没有共同的、公开的预期。再假设每个行为主体都在不停地观察市场，以此来发现有利可图的预期。接下来再进一步假设，每个行为主体在发现有利可图的预期时，就会采用它们，同时丢弃不如它们有利的其他预期。在这种情况下，行为主体的预期就是内生的，即独自适应于市场的当前状态。这样一来，他们意欲开发的市场，也就由他们自己来共同创造了。那么，市场将如何运行呢？它又将如何对资产定价呢？它会收敛到一个理性预期均衡吗？或者它会不会证实交易者所持的观点呢？

资产定价理论（**Asset**

Pricing Theory） 在资产市场上，假设行为主体是完全异质的，他们的预期需要不断适应市场，而市场本身则是他们的预期共同创造的。

在这里，我们提出这样一个资产定价理论。假设行为主体是完全异质的，他们的预期需要不断适应市场，而市场本身则是他们的预期所共同创造的。我们认为，在异质性假设下，预期将具有递归特性：行为主体必须根据对其他行为主体的预期进行预估来形成自己的预期，而且这种“自我参照”必定会将那种通过演绎推理形成的预期排除在外。因此，在无法通过演绎推理来判断其他人的预期的情况下，无论行为主体多么有理性，都将不得不对它们构造一些假说。因此，行为主体必须不断地形成一系列个人的、假说性的预期模型，或者说形成自己的“市场理论”模型，并对它们进行检验，然后根据预测得最好的模型进行交易。他们要时不时地舍弃“表现不佳”的假说，引入新的假说并检验。价格是由这些归纳出来的预期内生驱动的。因此，个人的预期要在市场中进化并“竞争”求存，而市场则形成于他人的预期。换句话说，各个行为主体的预期，是在它们共同创造的世界中共生进化的。

说到这里，一个很自然的问题是：这些异质性预期会不会进化成同质性的理性预期信念，从而证明有效市场理论？如果不是这样，那么会不会涌现出更加丰富多彩的个人行为和集体行为，从而证明交易者的观点，并解释上面提到过的各种现实中的市场现象？对于这个问题，我们将通过计算，而不是通过求解方程来回答。因为，我们的模型由于假设了完全异质性预期，所以太过于复杂，无法给出解析解。为了对我们这个内生预期市场的价格变化的动力学机制、投资策略和市场统计学特征进行分析，我们在一个基于计算机的市场，即圣塔菲研究所人工股票市场中，完成了一系列严格控制的实验。[5]

令人惊讶的是，我们通过计算机实验得到的市场图景，不但证实了经济学家的有效市场观点，同时也证实了交易者所持的观点。但是需要强调的是，这两种观点的证实是分别在两种不同的情况下，或者说分别在两种不同的市场体制下实现的。而且，在这两种情况下，我们开始实验的初始条件是一样的：都是赋予我们的交易者异质性信念，这些信念虽然彼此不同，但是都随机地聚集在一个接近同质性的理性预期均衡的

区间里。在实验中，我们发现，如果我们的行为主体根据观察到的市场行为来调整自己的预测的速度很慢，那么市场就会收敛到理性预期体制。在这种情况下，“突变”出来的预期，就无法获得能够盈利的立足点。此时，技术交易、泡沫、崩盘、自相关行为等现象不会出现。交易量也保持在低位。这就验证了有效市场理论。

如果我们使交易者以更快、也更加符合现实的速度适应观察到的市场行为，那么不同的信念将持续存在，而且市场将自组织成一个复杂的体制。在这种情况下，实验中就可以观察到丰富多彩的“市场心理”现象，即各种各样的预期。技术交易也成了一种有利可图的活动，市场中不时会出现暂时的泡沫和崩溃。另外，交易量很高，而且交投清淡时段和交投活跃时段交替出现。价格时间序列显示，波动是持续发生的，完全符合现实金融市场的价格序列的“广义自回归条件异方差行为”特征。高交易量也呈现出了持续性。在我们的实验中，投资者个人的行为会不断进化，而不是固定不变的。在这种情况下，交易者的观点被证明是合理的。

在接下来的内容中，我们将先讨论以内生预期方法来研究市场行为的理论基础，并引入条件预期假说或“预测器”集合概念。接下来，我们构建了一个计算模型，作为我们分析的基本框架，然后描述这个模型在计算机系统中实现的方法。接下来，我们要讨论计算机实验的结果，并对我们的发现与相关文献中运用其他现代方法得到的结果进行比较，并给出我们的结论。

为什么归纳推理是有效的

在进一步讨论之前，我们要先证明，一旦我们引入了异质性的行为主体，那么行为主体的演绎推理就会遭到失败。我们强调，在演绎推理不可行的情况下，行为主体必须诉诸于归纳推理。在金融市场中，这是很自然的事情，也是实际发生的事情。

演绎推理导致不确定的预期

在这里，我们使用一个简单的套利定价模型，来说明行为主体的演绎逻辑为什么是不确定的。在这里，我们暂不讨论这个模型的技术细节。因为，这个套利定价模型是后面资产市场模型的一个特例：只需假设风险系数λ任意接近于0，且假设预期符合高斯分布。考虑具有一个只

有一种证券的市场，该证券所提供的随机的收益报酬或股息序列为 $\{d_t\}$，而无风险的场外资本的收益则为每期固定的r。在每一期t，给定市场信息I_t，每个行为主体i都可以形成他个人对下一期的股息和价格的预期，分别为$E_i[d_{t+1}|I_t]$和$E_i[p_{t+1}|I_t]$，并令这些预期组合的条件方差为$\sigma^2_{i,t}$。假设存在完全套利，那么资本市场在如下均衡价格时出清：

$$p_t=\beta\sum_i w_{j,t}(E_j[d_{t+1}|I_t]+E_j[P_{t+1}|I_t]) \quad (1)$$

换句话说，该证券的价格p_t是在这样一个价值时被叫买的，它反映了各个行为主体的市场预期的当前（加权）平均值，并以贴现因子$\beta=1/(1+r)$贴现，其中权重 $w_{j,t}=(1/\sigma^2_{j,t})/\sum_K 1/\sigma^2_{k,t}$，是对行为主体$j$的预测的“相对置信度”。

再假设这些行为主体都是“聪明”的投资者。关键问题是，每个行为主体对股息的预期和$E_i[d_{t+1}|I_t]$和对价格的预期$E_i[p_{t+1}|I_t]$是如何形成的。解释行为主体可以理性地，即通过演绎逻辑方法，形成这些预期的标准理由如下。假设投资者都是同质的，它们（i）以同样的方式运用可得信息I_t来形成对股利的预期；以及（ii）知道其他行为主体都使用相同的预期；再进一步假设行为主体（iii）是完全理性的，无论怎么难的逻辑推理都能完成；（iv）知道每一期的价格都是套利形成的，如式（1）所示；以及（v）知道上面（iii）和（iv）这两点是共同知识。这样一来，根据定义，所有行为主体对未来股息的预期$E_i[d_{t+k}|I_t]$是已知的、共享的和相同的。而且，因为假设了同质性，所以我们可以删除行为主体下标，并将权重设为$1/N$。在此基础上，运用标准的证明方法不难证明，只要将式（1）用于未来$t+k$期，求出各期的预期，然后反复地向后替换$E[p_{t+k}|I_t]$，各行为主体就可以迭代方式求出当前价格p_t：[6]

$$p_t=\sum_{k=1}^{\infty}\beta^k E[d_{t+k}|I_t] \quad (2)$$

如果对股息的预期是无偏的，那么平均来说市场将支持对股息的预测，价格序列也将处于理性预期均衡。因此，当信息随着时间的流逝而波动时，价格也随之波动，并且将反映证券的“正确”价值或“基本”价值，从而使得投机行为无法不断地获得利润。当然，经济学文献中的理性预期模型通常比这里所说的更加复杂一些。到目前为止，如果我们愿

意采用上述假设，这些假设在很大程度上依赖于同质性假设，那么资产的定价便可以通过演绎方法来确定，这就是说，至少从原则上看，行为主体是能够推断出当前价格的。

现在引入一个更符合现实的假设，即假设交易者虽然都是“聪明”的，却是异质性的，每个行为主体都可能不同于其他行为主体。在这种假设下，可得的共享信息I_t包括了过去的价格、过去的股息、交易量、经济指标、小道消息、新闻，等等。不过，这些只是定性信息再加上一些数据序列，而且可以有许多种不同的、完全言之有理的统计方法用来预测未来股息，它们分别基于不同的假设和不同的误差标准。因此，并不存在“有客观基础的”、可以让不同行为主体用来实现彼此协调的预期模型，这也就意味着，没有任何“客观的”方法，可以让一个行为主体获悉其他行为主体对未来股息的预期。这个事实足以令式（1）中的资产价格产生不确定性。而且，更加糟糕的是，异质价格预期$E_i[p_{t+1}|I_t]$也是不确定的。这是因为，假设行为主体i尝试“理性地”演绎推理出这个预期，那么他就要根据式（1）求得$t+1$期的市场出清时的预期：

$$E_i[P_{t+1}|I_t]=\beta E_i\left[\sum_j\{w_{j,t+1}(E_j[d_{t+2}|I_t]+E_j[P_{t+2}|I_t])\}\middle|I_t\right] \quad (3)$$

而这就要求行为主体i在形成他对价格的预期时，必须考虑他对其他人在本期起两期内关于股息和价格，以及相对市场权重的预期的预期。类似地，要想消去另一个未知项，即价格预期$E_j[p_{t+2}|I_t]$，也需要进一步的迭代，但是这要求行为主体必须考虑他们对其他行为主体关于第$t+3$期的未来股息和价格的预期的预期的预期。这就像凯恩斯在1936年谈到过的，必须考虑“对平均意见的平均意见”的预期。

在假设了同质性的条件下，对他人的预期的预期将“坍塌”为单一的、共同的、客观决定的预期。但是，在假设了异质性的条件下，不仅无法通过任何客观手段去获悉他人对股息的预期，而且试图消除另一个未知项，即价格预期，也只会导致主观预期的主观预期的重复迭代，或者等价地对他人的主观先验的主观先验，这将导致“主观性”的无限倒退。此外，这种倒退还可能导致不稳定。如果投资者i认为，其他人认为未来价格会上涨，他可能会修正自己的预期，即预期价格会上涨。如果他认为，其他人认为价格可能会回到更低的位置，那么他也可能会修正自己的预期，即预期价格将回落。因此，我们不难想象，只要稍有风吹草动，有任何一点表明其他人对他人的信念的信念有所改变的迹象，或者想象中的迹象，投资者的信念就可能非常迅速地出现“波动”和“转

变”。

因此，在异质性条件下，演绎逻辑将导致不确定的预期。需要提请读者注意的是，要证明我们在这里阐述的论点，并不需要假设行为主体的推理能力是有限的。我们的观点无非是，给定行为主体预期的差异，没有什么逻辑手段可以形成客观的预期。因此，市场中的完全理性也就不能明确界定。即使拥有无限智能，行为主体也不能以某种确定的方式形成预期。

归纳推理形成预期均衡

如果异质行为主体不能以演绎推理方式形成自己的预期，那么他们到底怎样才能形成预期呢？他们可以观察市场数据，可以猜想市场行为和其他投资者的行为的性质，还可以通过复杂的主观推理得到预期模型。但是归根结底，所有这些模型都是而且只能是一些假说。要想验证它们，除了观察它们在市场中的表现之外，没有任何客观的方式。因此，我们的计算机实验中的行为主体所面对的如何选择适当的预测模型的问题，与统计学家在为特定数据集选择适当的预测模型时所面对的问题，在性质上是一样的。后者也没有客观的方法去决定应该选择什么样的函数形式。当然，由于投资者选择的预期模型还会影响价格序列，这就使得我们这里的情况更加复杂。这就意味着，如果把这些投资者当成“统计学家”的话，那么他们对模型的选择会影响数据，进而又会影响对模型的选择。

在这里，我们假设，每个行为主体都扮演着“市场统计学家”的角色。[7]每位“市场统计学家”都会不断地构建出多个“市场假说”来，它们都是主观的预期模型，即什么在影响着市场价格和股息的变动。而且，每个“市场统计学家”都会同时测试若干个这样的模型。其中一些模型对市场走势的预测比较准确，那么“市场统计学家”就会“信任”这些模型，将它们“保留”下来，并根据它们来进行买卖决策、采取行动。其他“表现”不好的模型则会被舍弃。此外，“市场统计学家”还会不时构建一些新的模型，并利用市场来检验它们的准确性。随着时间的流逝，哪些预期模型的预测比较准确会越来越清楚，同时预测不准的那些模型则会被更好的模型取代，这个过程也就是行为主体学习和适应的过程。我们把这种行为，即找到适当的假说模型来作为采取行动的依据，增强对那些被证明有效的模型的信心，并舍弃那些没有通过验证的模型称之为归纳推理。[8]当问题无法明确界定时，这种推理方法非常有效。我们把那些运

用归纳推理来进行决策的行为主体称之为“归纳理性主体”。[9]

每个在归纳的意义上理性（下文简称为“归纳理性”）的行为主体，都会在自己的头脑中生成很多个预期模型。这些预期模型要“竞争上岗”，然后根据各自的预测能力来决定它们是“生存下去”还是被改变或被替换。行为主体的假说和预期，都要进行调整以适应当前的价格和股息模式。同时，价格和股息模式也要发生变化，以反映行为主体的当前假说和预期。由此，我们立即可以看出，市场是有“心理”的。我们把它定义为市场假说集，也称为预期模型集或心智信念集。在给定时间内，这个集合是行为主体采取行动的依据。

如果在价格和预期的形成过程中存在一些吸引子，那么这种“市场心理”就可能会收敛到一个稳定的、不变的异质或同质信念。这样的集合在统计学上是可以验证的，因此可以构成一个理性预期均衡。我们对市场是否会收敛到这样的均衡进行了研究。

归纳预期的市场

模型

在本节中，我们构建一个简单的资产市场模型，它的思路与布雷（Bray）以及格罗斯曼（Grossman）和斯蒂格利茨（Stiglitz）一脉相承。从结构上看，这个模型是新古典的，但是它也偏离了标准模型，因为它假设了异质性行为主体，而且这些行为主体是通过上述过程形成预期的。

现考虑这样一个市场：N个异质性行为主体要确定自己最想要的资产组合，他们可以在一只股息随机的有风险的股票，与一只无风险债券之间进行选择。这些行为主体的预期是各自分别形成的，不过在其他方面则是相同的。他们的效用函数都是常数绝对风险厌恶型的（CARA），形式为$U(c) = -\exp(-\lambda c)$。他们相互之间没有交流，每个人都既不会把自己的预期告诉他人，也不会透露自己的买卖意图。时间是离散的，用t来表示，时间期限则是不确定的。无风险债券的供给是无限的，而且利率恒为r。股票每次发行N单位，支付的股息为d_t，遵循给定的外生随机过程$\{d_t\}$，但行为主体不知道这个过程。

在这个模型中，股息的过程是可以任意给定的。在我们已经完成的计算机实验中，我们将它指定为一个AR（1）过程，即一阶自回归过

程：

$$d_t = \overline{d} + \rho\,(d_{t-1} - \overline{d}) + e_t \tag{4}$$

其中e_t服从独立同分布的高斯分布，并且均值为0，方差为σ_e^2。

在每一个交易期，每个行为主体都在无风险资产和股票之间进行分配，试图实现资产组合最优化。再假设行为主体i在第t期下一期的价格和股息的预测服从正态分布，其条件均值为$E_{i,t}[p_{t+1}+d_{t+1}]$，方差为$\sigma^2_{t,i,p+d}$。现在，我们可以讨论这样的预期是如何“达至”的。众所周知，在假设了常数绝对风险型效用函数和高斯分布的条件下，行为主体i对于持有风险资产份额的需求$x_{i,t}$可以用下式给出：

$$x_{i,t} = \frac{E_{i,t}\,(p_{t+1} + d_{t+1} - p\,(1+r))}{\lambda\sigma^2_{i,t,p+d}} \tag{5}$$

其中p_t是风险资产在第t期的价格，λ是相对风险厌恶程度。

总需求必须等于发行的股份数量，即：

$$\sum_{i=1}^{N} x_{i,t} = N \tag{6}$$

有了这个式子，模型就完整了，出清价格p也就可以确定了，即上面的式（5）中的当前市场价格。

在市场上，如果能够搞清楚入市的时机，无疑是有益的。在期间t的开始阶段，当前的股息d_t会被公布出来，而且所有行为主体都可以观察到。然后，行为主体运用这个信息，以及关于市场状态的一般信息，去形成他们对下一期的价格和股息的预期$E_{i,t}[p_{t+1}+d_{t+1}]$。其中，“关于市场状态的一般信息”包括历史上的股息序列{……d_{t-2}，d_{t-1}，d}和价格序列{……p_{t-2}，p_{t-1}}。然后，他们就可以计算出他们想要持有的资产组合的各个参数值，并将他们的需求传递给专家，后者公布一个p_t以出清市场。在下一个期间的开始阶段，新的股息数d_{t+1}被公诸于众，同时第t期的各预测器的精度都将得到更新。以后各期依此类推，重复进行。

对预期形成建模

遗传算法（**Genetic**

Algorithm）　约翰·霍兰德提出的一种算法，通过模拟自然进化过程搜索最优解。

到目前为止，我们已经有了一个简单的、新古典双资产市场。接下来，我们要打破传统，允许我们的行为主体各自通过归纳推理形成自己的预期。要做到这一点，一个显而易见的途径是，先赋予每个行为主体一个属于他自己的预期模型，这些模型有共同的函数形式，只不过其参数的更新是因行为主体而异的，即从不同的先验参数开始，每个行为主体通过某种自己的方法，如最小二乘法进行更新。但是在这里，我们没有采用这种方法，而采用了另一种能够更好地反映前面所描述的归纳推理过程的方法。我们假设，每个行为主体在任何时刻都拥有很多个线性预测模型，即关于市场变动方向的假说或“市场理论”，并且会运用那些最适合当前的市场状态且最近被证明最可靠的模型。然后，行为主体会进行学习，但不是通过更新参数，而是通过发现他们自己的假说中哪些被“证明”是最好的，同时还会通过遗传算法不时地开发出新的模型来。这样一个模型有以下几个理想的性质：它能够避免因固定的、共同的函数形式而导致的偏差；它允许预期的“个性”随着时间的流逝而涌现出来，而不是简单地在某个先验信念的基础上构造出来的；它能够更好地反映实际的认知推理过程，即不同的行为主体可以“识别”出不同的模式，并且会从相同的市场数据中得出不同的预测。

在我们这个模型里，行为主体的预期是这样形成的。在每个期间，当前和过去的价格和股息的时间序列用一个J位的位串数组，可以理解为J个“市场描述器”给出的信息的集合，而行为主体的主观预期模型则用多组预测器来表示。每个预测器都是一个“条件-预测”规则，这有点类似于霍兰德所说的分类器，只不过后者是“条件-行动”规则。它由两部分组成：一是市场的当前状态可能会满足的市场条件，二是用来预测下个期间的价格和股息的公式。每个行为主体都拥有M个这样的预测器，他们心里同时会考虑M个关于市场的假说，并且会使用活跃的预测器，即与市场的当前状态相匹配的预测器当中最准确的一个，进行预测。这样一来，给定这些市场模式，每个行为主体都有能力“识别”出关于市场的不同状态集合，并给出适当的预测。

接下来，为了进一步阐明这个模型，我们简要地描述一下，这个预期系统在计算机上是怎样实现的。假设我们通过一个$J=13$位的位串数

组来总结市场的状态。例如，其中第五位可能对应于“价格在最近三个期间都在上升”，第10位则可能对应于“价格已经超过了股息与利率之比的16倍”。在这样的位串数组中，用“1”表示所描述的状态出现了，用“0”表示该状态不存在或未发生。在每个预测器中，条件部分对应于这些市场描述器，因此也是一个13位的位串数组，每个位置上都填上了“0”“1”或“#”。“#”的含义是“不关心”。如果预测器中的条件数组中所有的“0”或“1”，都与描述器的位串数组对应位置相同，同时所有“#”都对应于一个“0”或“1”，那么我们就说预测器的条件数组匹配了当前的市场状态，或者说“识别”出了当前的市场状态。举例来说，条件（####1########）“识别”出了“价格在最近三个期间都在上升”这种市场状态。条件（######### 0 ###）则“识别”出了“价格没有超过了股息与利率之比的16倍”这种市场状态。每个预测器的预测部分则是一个参数数组，它会触发一个与之对应的预测表达式。在我们的实验中，所有预测器都使用了价格和股息的线性组合$E[p_{t+1}+d_{t+1}]=a(p_t+d_t)+b$。这就是说，每个预测器都保存了$a$和$b$的一组特定的值。因此，一个完整的预测器（#### 1 #### 0 ###）/（0.96，0）就可以解释为“如果价格在过去三个期间内一直在上升，且如果价格不超过股息除以利率r的16倍，那么就预测下一个期间的价格加股息之和相当于本期间的96%”。这个预测器能够识别出市场状态（0110100100011），或者说将被该市场状态激活，但是却不会对市场状态（0110111011001）做出反应。

很显然，在能够识别出多种市场状态的那些预测器当中，只有很少的“1”和“0”。那些更“专门化”的预测器则有更多的“1”和“0”。在实际的计算机实验中，我们为每个行为主体准备了一个全部由“#”组成的默认预测器。然后运用遗传算法创建新的预测器，途径有两个：要么让预测器的位串数组中的值发生“突变”，要么通过对一个预测器的位串数组的一个部分，与另一个预测器的位串数组的互补部分进行“重组”。

在每一个时期，每个行为主体观察市场的当前状态，并且注意到他的预测器中的哪一个与该状态相匹配。这也就是预测系统的工作方式。行为主体会对所有活动的预测器当中最准确的H个预测器给出的线性预测进行“统计分析”，预测下一个期间的价格和股息，然后根据得出的预期值及其方差，运用式（5）计算出想持有的股票数量，并生成适当的出价或要价。一旦市场出清了，下一个期间的价格和股息就会公布出来，同时活跃的预测器的准确性记录也会得到更新。

如上所述，在这个预期系统中，学习是以两种方式进行的。第一种学习是快速的，出现在当行为主体知道哪些预测器是准确的，且可以作为自己采取行动的依据，哪些预测器是应该被忽略的这种情况下。另一种学习则比较慢，出现在当遗传算法需要不时地丢弃不好的预测器，并创建新的预测器的时候。当然，这些新的、未经检验的预测器不会造成破坏，因为只有当它们被证明准确时，它们才会被作为采取行动的依据。这就避免了脆弱性，而且保证了机器学习理论家所称的学习过程的“优雅性”。

现在，我们应该可以看出，这种多位、多预测器架构有四个优点。

第一，这种预期架构允许市场在各种不同的状态或情况下，出现各种可能的动态变化，即呈现出不同的特性。因为每个预测器都是一个模式识别型的预期模型，因此可以“识别”出这些不同的状态，行为主体可以“记起”给定的某种状态下，以往曾经发生过的事情并激活适当的预测。这就使得行为主体能够在市场变化时，迅速地转换预测行为，就像迅速的“态度转变”一样。

第二，这种架构能够避免因特定的关于预期的函数形式而导致的选择偏差。虽然我们的预测器的预测部分是线性的，但是在任何一个时间，以市场条件的多种组合为条件的预测器，其多样性都是有保证的。并且对于任何一个行为主体来说，预测表达式也总是非线性的。从形式上看，它是一个分段线性的、非连续预测函数，其域为市场状态空间，并且其准确度适应于该空间的不同区域。当然，由于用“二进制”的描述器来描述市场条件，这肯定会使预测受到一定的限制。

第三，学习集中发生在需要学习的地方。例如，$J=12$位的描述器就可以区分4 000多种不同状态的市场状态。然而，在所有这些状态中，只有少数几种会经常发生。用于识别不经常出现市场状态的那些预测器条件，不会被很频繁地使用，它们的准确性也不会被很频繁地更新。并且在其他条件相同的情况下，它们的精度也将较低。因此，它们在预测器之间的竞争中，生存下来的可能性很小。这样一来，预测器将聚类在更常被“访问”的那部分市场状态空间上，而这正是我们所希望看到的。

第四，也是最后一个优点，对于描述器的位串，可以进行分类或将它们组织成若干信息集，以便反映各种基本面因素（如价格-股息比）和技术面因素（如价格变动趋势等技术交易指标）。这种设计使我们能够准确地跟踪，哪些信息，即描述器中的哪几位，是行为主体正在使用或

被忽略的。如果我们想研究技术交易是如何“涌现”出来的，这一点至关重要。在这样一个模型中，我们还可以设置不同的行为主体“类型”，即根据所访问的信息集的不同，来对行为主体进行分类。不过在本章中，所有行为主体都能“平等地”观察到所有市场信息。

在上述优点中，有几个或许也可以通过神经网络来实现。但是，神经网络不如我们的预测器系统透明。我们可以很方便地对预测器系统加以监视，以观察在每个时期，哪些信息被行为主体单独地或共同地使用。

两种市场体制的涌现

实验设计

我们将用一系列实验从计算的角度研究我们的内生预期市场的行为。我们在所有这些实验中都保持了相同的模型参数，这使得我们可以利用这个模型，在相同条件下通过只允许受控的变化来比较市场结果。每个实验都运行25万个周期，来允许渐近行为出现（如果存在的话）；并在不同的随机“种子”下运行25次，以便收集截面统计数据。

我们所用的基本模型如上节所述。在实验中，我们通过选择参数值，以及在必要时选择函数形式来运行特定的模型。具体参数值如下：$N=25$个行为主体，每个行为主体有$M=100$个预测器，这些预测器都是条件$J=12$位的市场描述器。股息服从如式（4）所示的一阶自回归过程，其自回归参数ρ设定为0.95，这是一个接近于随机游走、但持续的过程。

反映市场状态下的12位二进制描述器的含义如下：

1～6：当前价格×利率/股息＞0.25，0.5，0.75，0.875，1.0，1.125

7～10：当前价格＞过去5期，10期，100期，500期的移动平均价格

11：始终打开（“1”）

12：始终关闭（“0”）

前6个二进制描述符，即前6位反映了当前价格与当前股息之间的关系，用来说明股票的当前价格是高于还是低于其基本价值的。我们把它们称为“基本面”位。第7～10位则是“技术面”位或“技术交易”位，用来说明是不是存在某种价格运行趋势。如果没有用处的话，第7～10位将被忽略；如果出现了某种技术分析趋势，那么它们就会被采用。最后两位一直固定为“0”或“1”，用来对实验进行控制。这两位并不传达有用的市场信息，但可以告诉我们，行为主体会在任何时候在何种程度上根据无用信息采取行动。当某一位上是“0”或“1”时，我们就说这一位是“配置好”的。此外，预测器会被随机选择来用于重组，而且，当其他方面相同时，特异性较高的预测器，即它们包含的“配置好”的位数越多，被选择的概率就越低。这样一来，就引入了一种趋向“全#配置”的微弱的漂移机制，从而可以确保由特定比特位表示的信息，仅在行为主体发现它在预测中真正有用时才会被使用。有了这种市场信息机制，我们就可以讨论“涌现”了。例如，我们可以说，如果第7～10位在统计上明显更多地被“配置”了，那么就表明，市场中已经涌现出了技术交易模型。

我们假设，预测是这样形成的：每个预测器j存储价格和股息的线性组合$E[p_{t+1}+d_{t+1}]=a(p_t+d_t)+b$中的参数$a$，$b$的值，同时，每个预测器也会将它自己的预测方差的当前估计值保存下来。

在进行计算机实验之前，我们在基于计算机的模型上进行了两个测试用的预实验。第一，我们测试了这个模型是否可以复制标准理论的理性预期均衡（r.e.e）。我们先用解析方式，计算出同质理性预期均衡（h.r.e.e）下的预测参数a和b，然后将所有预测器的参数值固定为这些计算出来的结果进行测试。我们发现，这种预测确实是成立的：模型复制了同质理性预期均衡，它保证我们的计算机化模型能够正确地工作，这体现在它的预期、需求函数、行为聚合、市场出清和时间序列上。在第二个测试中，我们将一个给定的股息序列，以及一个与之对应的事先计算出来的同质理性预期均衡价格呈现给各个行为主体，并检验他们是否能够独立地学会掌握正确的预测参数。他们确实能，尽管有一点点变化，这是因为他们一直在探索预期空间，这就保证了我们的行为主体一直正确地学习。

实验

我们用基于计算机模型完成了两组基本实验，它们分别对应于行为主体对替代预期的“慢速探索”和“适中速度探索”。这两组实验产生了两

种不同的市场体制，即市场的两种不同的特征性行为。在慢学习速度的实验中，遗传算法平均每1 000期才被调用一次，预测器发生转换的概率为0.3，预测值的精度更新参数θ设置为1/150。在适中探索速度的实验中，平均每250期调用一次遗传算法，预测器发生转换的概率为0.1，预测器的精度更新参数θ设置为1/75。[10]另一种做法是，在两组实验中保持相同的模型参数，然后将一些随机选取的预期参数，从以计算出来的同质理性预期均衡为中心的均匀分布中选取出来，在实验开始时赋予行为主体。

在“慢速探索”的实验中，没有任何非理性预期均衡的预期可以获得一个稳固的立足点：市场进入了进化稳定的理性预期均衡。而在“适中探索速度”的实验中，我们却发现，市场进入了一个复杂的状况：“市场心理”驱动的行为出现了，市场行为显著地偏离了理性预期均衡的基准，真实金融市场在统计意义上的“基本特征”也都被观察到了。

接下来，我们详细描述这两组实验，以及它们归纳出的两种市场体制或“相”。

理性预期体制

如上所述，在这组实验中，行为主体在预测空间中不断探索，但是探索的速度很低。我们在这些实验中观察到，市场价格迅速收敛于风险调整后的同质理性预期值，即使行为主体以非理性预期开始，也是如此。换句话说，在这种情况下，对于一个具有内生的、归纳预期的市场来说，同质理性预期是一个吸引子。[11]当然，这个结果并不令人惊奇。即便这些行为主体的预测不同于同质理性预期均衡值，那么大多数其他行为主体的预测都接近于同质理性预期均衡这个事实也足以保证，市场出清价格能够纠正这些偏离的预期。这就是说，有一个自然的（虽然比较弱）趋向同质理性预期均衡的吸引子。

不过，在这种体制下的均衡与标准理论中的理性预期均衡，有两个方面的不同。第一，在这种体制下，不假设演绎方法可以带来均衡，事实上均衡也不是通过演绎方法得到的。我们的行为主体是以归纳方法实现同质理性预期的，它与标准理论的同质理性预期只有部分重叠。第二，我们这里的均衡是随机的。行为主体不断地探索各种替代品，尽管探索的速度很低。这种替代探测，尽管规模可能很小，但是确实在系统中引入了一些“热噪声”。正如我们所期待的，在这种体制下，行为主体的持股量保持了高度的同质性，交易量却依然很低（仅仅反映了由于突

变和重组导致预测的变化），同时泡沫、崩溃和技术交易等现象也都不会出现。我们可以说，在这种体制下，有效市场理论及其政策含义依然成立。

复杂体制

我们现在允许行为主体在信念空间中，以更加符合现实的速度进行探索。在这组实验中（如图3-1所示），价格序列几乎与理性预期体制中的价格序列完全相同。由于“探索活动”大为增加，导致了更高的方差，这种风险使得价格更低。

然而，仔细检查了实验结果之后，我们发现一些复杂的模式已经在信念集中涌现出来了，而且此时的市场显示出与理性预期体制下的市场非常不同的特征。例如，当我们放大两个价格序列之间的差异时，我们就可以观察到暂时的价格泡沫和崩溃的系统性证据（如图3-2所示）。我们把这一系列新的市场行为称为复杂体制或丰富的市场心理体制。

出现了泡沫和崩溃现象这个事实表明，市场上已经涌现了技术交易这种形式，即根据趋势买入或卖出。我们可以通过考察行为主体，据此做出预测的信息来检验这一点。图3-3给出了随着时间的推移，全部预测器当中被使用的“技术交易位”，即为“1”或“0”的数量。在这两组实验中，我们在设置初始状态时，都在预测器中随机地植入了技术交易位。计算机实验结果表明，在理性预期体制下，技术交易位并未能提供有用的信息，并且后来都变成了无用的预测器而被舍弃了。而在复杂体制下，它们在预测器“种群”中自我繁殖，然后在经过了大约15万期之后达到了稳定状态。这就是说，在复杂体制下，技术交易一旦涌现出来，就能够维持下去。[12]

复杂体制下价格的统计特征，明显不同于理性预期体制下。这主要是因为在复杂体制下，“峰态”非常突出（如表3-1所示），被交易换手的股票数量按每1万期计，在复杂体制下要比理性预期体制下高出大约300%。这说明，尽管市场一直在进化，但是行为主体的预期的异质性仍然一直保持着。正如我们所知道的，厚尾性和大交易量也是实际金融市场的价格数据的典型特征。

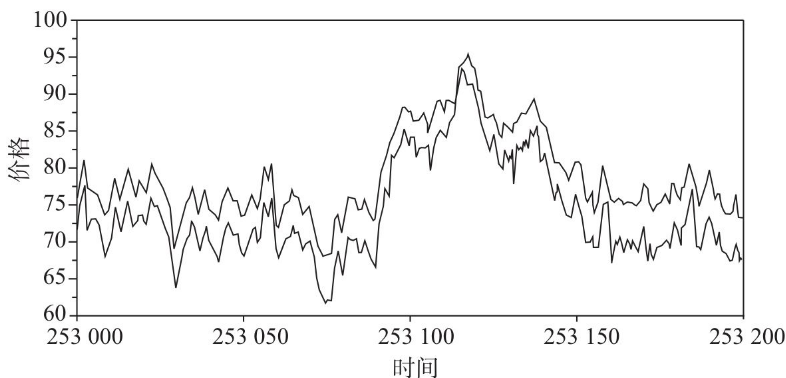


图3-1 理性均衡价格与“复杂体制”下的价格对比

图中的两个价格序列是在相同的随机股息序列之上生成的。上面的是同质理性预期均衡价格，下面的是复杂体制下的价格。后一种情况下方差较高，由于行为主体的风险规避，这导致了较低的价格。

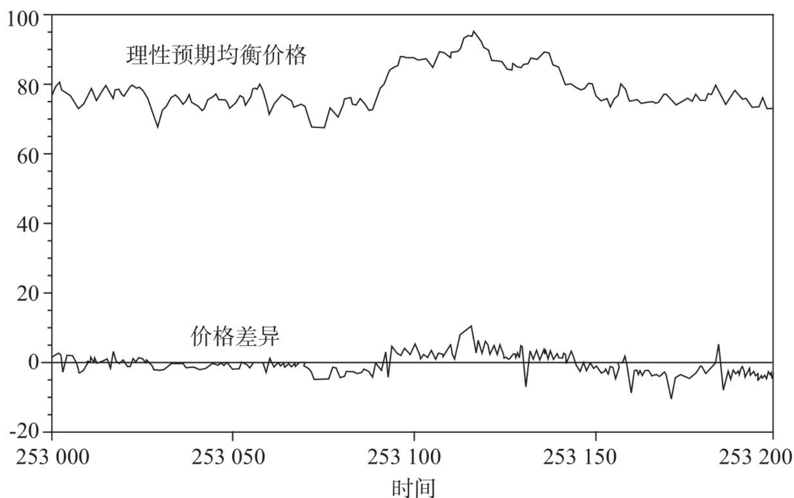


图3-2 复杂体制下的价格对基本价值的偏离

图的下部给出了图3-1中的两个价格序列之间的差异，上部给出的是同质预期均衡价格序列。为了便于观察和比较，对复杂体制下的价格序列进行了缩放，以匹配理性预期均衡价格序列，而且两个价格序列之间的差异也放大了一倍。

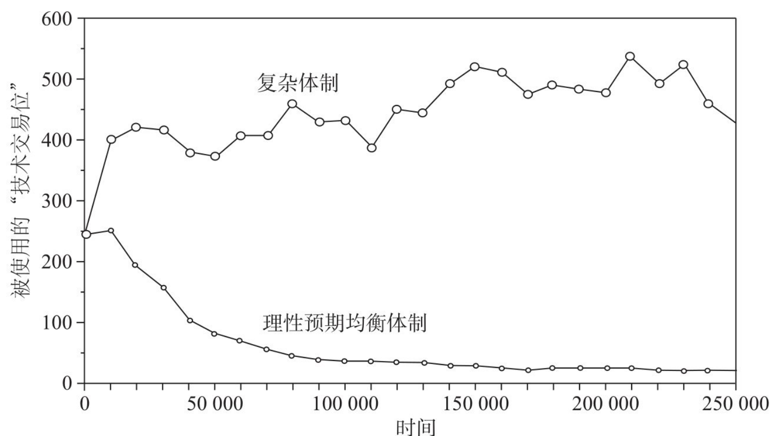


图3-3 市场进化中被配置的“技术交易位”的数量

图中显示的是在两个体制下进行的25个实验所得到的数据的中位数。

那么，在“丰富多彩的心理体制”或复杂体制下，技术交易是如何涌现的？答案是，在这种情况下，探索的“热度”足够高，能够抵消趋向理性预期均衡的“天然”吸引力。因此，非理性均衡预期的信念子集并不会迅速消失。事实上恰恰相反，它们可以相互强化。例如，假设有些预测器在比较早的时候，出于偶然，对当前呈现上升趋势的市场给出了一个价格将会上涨的预测，那么拥有这些预测器的那些行为主体，就更有可能在上升趋势中看高市场、买入股票，从而使价格高于原本可能会出现的价格，这就会导致轻微的向上偏差。这种偏差足以证明这些规则的有效性，并保证它们可以留在市场中。对于那些预测价格将回归到基本价值的预测器，类似的事情同样会发生。当然，这样的预测器出现的“密度”必须足够高，这样才能验证彼此的有效性，从而保证大家都能在预测器“种群”中存在下去。这里发生的情况可以用生命起源理论所描述的情况来类比。根据生命起源理论，在由单体和聚合物组成的生命原“汤”中，相互增强的RNA单元的密度必须足够高，不然这些会复制的单元就无法获得稳固的立足点。

表3-1 收益和成交量统计（中位数）

	均值	标准偏差	偏度	峰度[13]	成交量
理性预期均衡体制	0.000	2.1002	0.0131	0.0497	2 460.9
复杂体制	0.000	2.1007	0.0204	0.3429	7 783.8

注：在两种体制下，进行了超过25万期的实验后得到的数据。

13 这里的峰度指的是过量峰度，如峰度－3。

因此，如果趋势跟随或均值回归信念偶然地出现了，同时如果股息序列中的随机扰动激活了它们，并在随后验证了它们的有效性，那么技术交易就会涌现。而且，从涌现的那一刻起，它们就可以在行为主体识别出来的模式的“种群”中取得它们的位置，并且相互强化，持续地存在下去。从系统的子组件的相互作用中涌现出来的这种结构，证明我们将该体制称为“复杂体制”是正确的。

对于这种相互强化的预测子群体，至关重要的前提是，是不是存在作为它们出现条件的市场信息。市场状态就像“太阳黑子”一样，可以发出信号，使得相关预测器能够在它们与该信号相关联的方向上实现协调。当然，这些市场状态并非是那些不传递任何真实信息的太阳黑子可比的。这种协调或相互作用偶然建立起来之后，就有可能在市场中存留下去。我们可以说，市场状态可以成为信号这种能力，启动了最终导致复杂行为的相互性。而且，没有必要为了论证这一点，而假设一个单独的噪声交易者类别。在进一步的实验中，我们检验这个“信号猜想”，方法是把所有预测器中的条件部分都“关闭”起来，只需要将它们用不可替换的“#”填充满即可。现在，预测器无法区分各种不同的市场状态，所以市场状态就不能成为信号。实验结果与我们的猜想，即信号驱动我们所观察到的模式一致：复杂的体制没有形成。作为对技术交易信号的重要性的进一步测试，我们对当前价格与前期的技术指标的关系进行了回归分析（价格＞　500期移动平均线）。在理性预期体制下，技术指标完全不显著，而在复杂体制下，趋势指标是显著的。比如说，t值为5.1，是25个实验的样本的平均值。这就表明，这个指标确实携带了有用的市场信息。对实际的金融市场进行数据分析，也得到了类似的结果。

现实世界的金融市场有一个显著的特点，那就是：市场上的价格波动和交易量都会表现出持久性或自相关性，并且波动性和交易之间也存在显著的互相关。换句话说，在时长比较随意的各个期间内，交易量和波动率会保持在或高或低的水平上，而且两者之间是相关的。在我们的“归纳推理市场”中的复杂体制下，波动也呈现出了持续性的或波动性的“广义自回归条件异方差行为”的特征（如图3-4所示）：在恩格尔广义自回归条件异方差检验中，卡方检验结果为在95%水平上显著。[13]它还显示出了大交易量的持续性（如图3-5所示），以及交易量和波动率之间的显著互相关性（如图3-6所示）。在这里给出的图表中，与通常作为股票市场标准的IBM公司股票之间的相关性，也体现得非常清楚。当然，需要注意的是，因为我们的人工股票市场上的交易时段和实际市场的交易日并不完全匹配，我们不应该预期会看到百分之百的重叠。但是，从

定性的角度来看，我们的市场的持续性与IBM公司股票的持续性是完全类似的。这些相关性是标准模型没有解释的，因为在标准理论中，它们均为0。

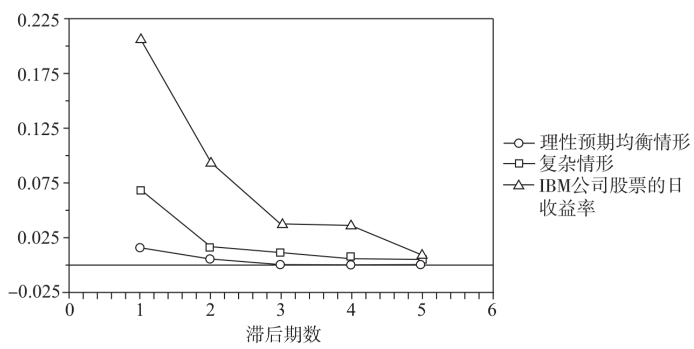


图3-4　三种情形下波动率的自相关性

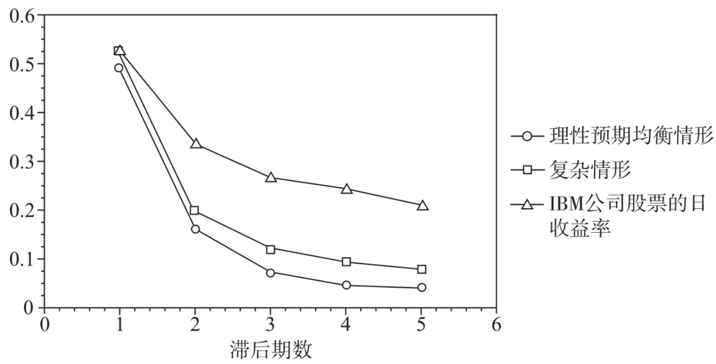


图3-5　三种情形下成交量的自相关性

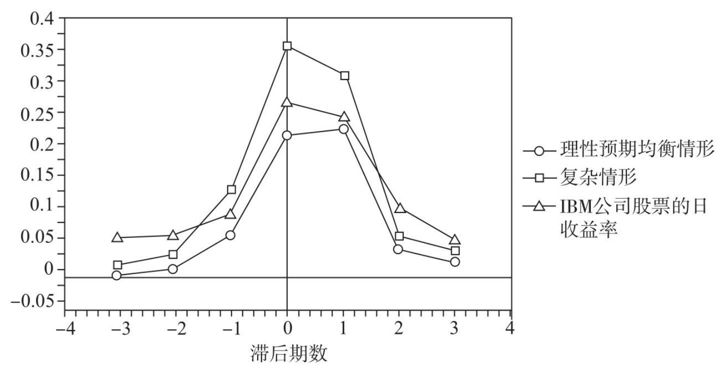


图3-6　三种情形下成交量与波动性之间的互相关性

为什么现实世界中的金融市场，还有我们的“归纳推理市场”，会呈现出这些特征，这仍然是一个悬而未决的问题。我们提出了一个简单的

推测性的进化论解释。无论是在实际市场，还是在人工模拟市场中，行为主体都在不断探索和测试新的预期。偶尔的、随机的、更成功的预期会被发现。这种预期将改变市场，并触发预期的进一步进化，使得各种大小“雪崩”式的变化，以级联放大的方式传遍整个系统。当然，在这个非常短暂时滞的尺度上，这些雪崩不是通过遗传算法传播的，而是以行为主体改变自己的活跃预测器的方式传播的。

然后，变化会加剧：波动率增大、成交量也增加。要检验这个推测的正确性，一种方法是考察自相关程度会不会随着式（1）中的预测器的精度更新参数θ的增大而增加。这是因为，θ值越大，个体行为主体在他的预测器之间进行“切换”的速度就越快，这种切换导致的级联放大也越显著。实验结果证明，自相关程度确实会随θ增加。这种级联性的切换有时会被市场“消化”而消失。因此，在我们的人工股票市场上，会出现“湍流时期”和“静止时期”交替出现的现象，这与实际市场一样。[14]

是普遍规律，还是人为假象

这种复杂体制的存在，会不会是我们设计模型时给出的假设所导致的一个人为假象？当然不是。我们改变了模型参数和预期学习机制，进行了稳健性检验，结果表明，复杂体制及相关的定性现象都是非常稳健的。它们不是模型中的某种缺陷带来的人为假象。[15]

有人可能会提出反对意见说，如果某些行为主体能够发现一种更加优越的预测手段来利用市场，那么就可以针对前述复杂模式进行套利，从而导致市场再次收敛到理性预期。我们并不这样认为。如果真的存在一个“特别聪明”的、可以利用他人预期的“元预期”模型，那么这样一个模型也必定是当前市场信息的一个复杂的非线性函数，因为它要聚合他人的预期。我们所假设的分段线性形式已经覆盖到了以当前市场信息为条件的非线性预期模型的空间。在这种假设下，行为主体其实已经通过遗传算法“取到”了这种“元预期”模型的近似形式。因此，复杂体制的存在本身并不构成对预测的限制，恰恰相反，在我们的内生预期模型中，市场信息事实上已经被当成了信号使用，从而打开了更广阔的可能性空间。特别是，市场可以通过自组织形成相互支持的预测器子“种群”。布洛克和霍默斯发现，在一个只存在几类内生信念适应交易者的更简单的解析模型中，也存在类似的丰富多彩的资产价格动态机制。没有理由认为，这些子群体涌现出来之后必然会处于随机均衡。行为主体会不断地

进行调整，以便使他们的预期相互适应，从而使得市场探索的可能性空间更大，因此市场是不稳定的。在一些早期的探索性实验中，我们“冻结”了成功的行为主体的预期，然后在更晚的时间重新注入这些行为主体，以及他们以前的成功预期。结果显示，这些重新被引入的行为主体的成功程度达不到平均水平，这就表明市场已经进化了而且不是稳定的。

也许，还有人会反对说，既然我们在预测器中使用了那些条件位，那么我们已经在模型中内置了技术交易。因此，技术交易会出现在复杂体制中，一点也不奇怪。但实际上，我们的模型内置的只是技术交易的可能性，而不是它的使用。在模型中，市场描述器的使用是禁止选择的。市场信号必须具有价值才会被使用，而且技术交易之所以会出现，完全是因为这些市场信号诱发了以这些市场信号为条件的、相互支持的预期。

在我们的模型中，有一个得到了明确界定的“市场心理”，它会不会也能体验到各种“情绪”呢？显然不能。但是，请读者注意，我们假设行为主体能接受各种关于市场的假说。我们完全可以想象这样一种情形：在长期“牛市”上升趋势下，市场价格远远高于基本价值，市场状态不但激活了继续指示上升趋势的预测器，同时也激活了预测价格将快速下跌而去修正的其他预测器。这种令人紧张的组合，在现实世界的金融市场中，在我们的人工股票市场中，都经常会出现。

那么，我们这个市场上的交易是如何发生的？行为主体进行交易的动机何在？理性预期理论认为，演绎理性行为主体没有交易的动机，即使他们的信念有所不同时也是如此。假设其他行为主体可以利用不同的信息集，每个行为主体最终都会“达至”相同的信念，这都是预设的。相比之下，我们的归纳理性行为主体（他们不能直接进行交流）的信念则不一定会收敛。因此，他们保留了交易的动机：根据自己作为市场统计学家的能力来“赌一把”。从表现上来看，我们的行为主体作为“统计家”拥有同等的能力，因此他们进行贸易是不合理的。但是，他们在这方面的能力是一样的，他们找到好的预测器的运气则会随时间流逝而变得发散。在每个时期，他们的预测器的准确性，充分体现在他们对无风险资产和风险资产之间的配置当中。给定行为主体的条件，只能把他们作为“市场统计家”来采取行动，他们进行交易的行为是符合理性的。

我们的内生预期理论，还与另外两种现代理论相契合。我们的模型推广了布雷等人的学习模型，他们的模型也假设了内生的预期更新。但

是，布雷等人的模型所假设的更新，是从一个共同的非理性预测进行同质性更新。而我们的模型则假设了异质性的行为主体，他们能够形成可以利用的以任何模式存在的预期。我们所采用的进化论方法也与布卢姆和伊斯利的进化模型有很强的“亲和力”。我们的模型和他们的模型，都假设了一个由一系列的预期规则，或者更准确地说是投资规则组成的“种群”，这些规则在市场中相互竞争以求得生存，并且不时地做出调整。但是在他们的模型中，关注的焦点在于相互竞争的、不同规则类型的选择性生存，而不是导致复杂现象产生的相互支持的“子种群”的涌现，也不是市场信号在这种涌现过程中的作用。

当然，我们的归纳理性市场，也抽象掉了现实世界中的金融市场的许多细节。在现实世界的金融市场中，投资者不能完全优化自己的投资组合，也不能在每个交易期都实现完全的市场出清。事实上，除了预期的形成这一部分，我们这个市场是简单的、新古典主义的。然而，我们的目标并不是构建一个“现实主义”的市场。相反，我们的目标是证明，当存在异质性时，丰富多彩的“市场心理”现象或行为的涌现是不可避免的，即便在新古典条件下也是如此。我们不需要像在其他研究中那样，假设共同信息来引出这些现象，也不需要假设共享预期和羊群效应。我们同样不需要援引“行为主义”，或其他任何形式的非理性。羊群效应和“准”理性行为，确实可能存在于实际市场中，但是它们并不是我们这些发现的必要条件。

金融市场是个复杂系统

在资本市场中，行为主体的预测创造了行为主体试图预测的世界。因此，资本市场是具有“反身性”的：价格是由交易者的预期产生的，同时这些预期又是基于对他人预期的预期而形成的。[16]预期的这种反身性或“自我指涉性”，排除了通过演绎推理形成预期的可能性，从而使得完全理性无法被明确界定。因此，行为主体只能把他们的预期作为一种假说：他们根据归纳推理结果来采取行动，并形成各种各样的个人的预期模型，这些是他们不断引入、检验、据以行动或者舍弃的模型。市场受预期的驱动，而预期则内生地适应这些预期所共同创造的生态。

我们在计算机上创造了一个人工股票市场，对这种内生预期市场模型进行了实验，得到的结果可以解释金融领域非常引人注目的一个难题：标准理论倾向于认为，市场是高效的，羊群效应是没有任何理由会

出现的，系统的投机利润也是不可能存在的；但是交易者则倾向于认为，市场会呈现出某种“市场心理”，从众效应明显是存在的，获得投机利润的机会也不会太少。最近的一些研究通过引入一些“行为”假设，如噪声交易商的存在，证明了交易者观点的合理性。但是，我们能证明，不需要引入“行为”假设，这两种观点都可以是正确的。归纳理性交易者市场可以存在于两种不同的体制下：在可替代性预测的探索速度足够低的情况下，市场会进入一个稳定的简单体制，它对应于有效市场理论中的理性预期均衡。在可替代性预测的探索速度更高、更符合现实的情况下，市场会自我组织成一个复杂体制，丰富多彩的“市场心理”行为将涌现，技术交易将出现，短暂的泡沫和崩溃也会出现。在复杂体制下，价格数据的统计特征，特别是“广义自回归条件异方差行为”，与实际市场数据如出一辙。在由各种预期组成的“种群”中，当涉及趋势追随或均值回归的个人预期相互强化时，当市场指标被用作协调这些相互强化信念的信号装置时，上面这些现象就会涌现出来。

我们的内生预期市场表明，在实际的金融市场中，如果存在归纳推理交易者，那么信念的异质性、对基本交易的偏离，以及时间序列的持续性，所有这些都可以无限期地持续存在下去。我们认为，实际的金融市场属于复杂体制。

COMPLEXITY

AND THE

ECONOMY

04

收益递增和路径依赖

技术竞争、正反馈及历史事件导致的锁定

复杂技术被采用后，往往表现出收益递增的特性。技术被采用得越多，获得的经验就越多，技术被改进得也就越多。当多种收益递增的技术为获得由采用者组成的市场而“相互竞争”时，某些看似微不足道的偶然事件可能变得相当重要，导致经济最终可能被“锁定”到某些不可预测的、较差的结果上。

收益递增（**Increasing Returns**） 是指这样一种趋势：如果一种技术处于领先地位，那么在“正反馈”机制的作用下，它就会拥有更进一步的优势，从

而可以获得更加领先的位置。

本章介绍收益递增经济学的基本概念：技术竞争、历史小事件导致的锁定，以及不可预测的、路径依赖的结果的可能性。我们早就知道，收益递增可能导致多重均衡，而且不是所有结果都是最优的。我们不知道的是，在多种可能的结果当中，被选中的那个特定结果是如何选择出来的。因为从静态的角度来看，哪个结果将被“选中”是不确定的。本章将递增收益问题重新定义为一个服从于特定随机事件的动力学问题，为上述选择问题提供了一个解决方案。在某个特定的小的随机事件集合下实现的“特定历史”中，某个结果被选择了。而在历史的另一种“实现”中，在另一个不同的小的随机事件集合下，被选择的却是另一个结果。因此，收益递增问题应作为随时间流逝而进化的概率系统来研究。这篇文章问世以后，这种方法逐渐成为了研究收益递增问题的公认的方法。

下面这篇文章最早是在1983年9月公之于众的，当时它是国际应用系统分析研究所的一篇工作论文。但是，由于它强调可能“锁定”到某些不可预测的、较差的结果上，它的发表之路异常艰辛。最后，直到1989年，这篇文章才发表在了《经济学杂志》（*Economic Journal*）上。不过自那之后，它就成了经济学中被引用最多的文章之一。[1]1994年，由美国密歇根大学出版社出版的我的文集《收益递增与经济中的路径依赖性》收入了这篇文章。但是，考虑到它是复杂经济学中一篇奠基性文章，我再次将它收录在本书中。

* * *

这篇文章探讨收益递增情况下分配的动力学。它考虑的是一个收益递增情况下自然会出现的情境：行为主体在相互竞争、以求被采用的技术之间进行选择。

现代而又复杂的技术被采用后，往往表现出收益递增的特性。技术被采用得越多，获得的经验就越多，技术被改进得也就越多。[2]当两种或多种收益递增的技术为获得由采用者组成的市场而“相互竞争”时，某些不重要的事件，可能会非常偶然地赋予其中一种技术被采用的初步优势。于是这种技术就可能比其他技术得到更好的改进，从而可能吸引更大比例的潜在采用者。因此，这种技术就有可能进一步被采用和被改

进。这样一来，某种在早期的采用中偶然地获得了领先优势的技术，就有可能最终占据由潜在采用者组成的全部“可能市场”，而其他技术则被锁定在该市场之外。当然，在不同的“微不足道的事件”发生时，如原型机获得了预料之外的成功、早期开发商的想法、政治环境的改变等等，被广泛采用、得到持续改进、最终主导市场的技术可能不同。技术之间的竞争有多种潜在的结果。

> 路径依赖（**Path Dependence**或**Path Dependency**） 指人类社会中的技术进化或制度变迁均类似于物理学中的惯性，一旦进入某一路径，就可能对这种路径产生依赖。

众所周知，在收益递增情况下，配置问题往往会表现出多重均衡性。因此，当出现多个结果时，我们本就不应该觉得奇怪。静态分析通常可以找到这些多重均衡，但是通常不能告诉我们哪一个均衡将被“选中”，动态方法则可能告诉我们更多东西。动力学分析允许技术被采用期间发生“随机事件”的可能性，因此通过动力学分析可以研究这种随机事件是如何影响对结果的“选择”的，即一系列随机的“历史事件”是怎样累积起来，并推动了一个趋向某种结果（即特定市场份额）的过程。同时，别的一系列随机的“历史事件”也会累积起来，并推动另一个趋向另一种结果（即另一种市场份额）的过程。这种分析还可以揭示人们熟知的收益递增的两大属性，即不可预测性和潜在低效率性是如何出现的，或者说收益递增是如何放大偶然事件在技术采用过程中的作用，以至于即使拥有关于采用者的偏好和技术发展可能性的先验知识，也不足以预测“市场结果”。还有，长期潜力较劣的技术，是如何在收益递增驱动下的技术采用过程中被开发出来的。这种动力学方法还指向两个新的性质：第一是刚性或无弹性，即一旦某种结果（主导技术）开始出现，它就会变得越来越被“锁定”；第二是非遍历性或路径依赖，即因为历史“小事件”并不会在动态过程中被“平均化掉”，也不会被“遗忘”，它们可能会决定结果。

本章对比了收益递增、收益递减和收益不变这三种情况下，技术

的“市场份额”的动态变化。它特别关注收益如何影响可预测性、效率、灵活性和遍历性，它也特别关注经济是怎样被“历史事件”锁定在劣等技术占垄断地位的境况中去的。

简单模型：仅存在两种技术

核能发电，可以采用轻水、气冷、重水或钠冷却反应堆。太阳能的利用，可以通过晶体硅或非晶体硅技术。我从大量诸如此类的实例中抽象出了一个初步的、简单的模型。在这个模型中，我假设存在两种新技术A和B，它们为了让自己被大量的经济行为主体采用而“相互竞争”。这两种技术不是任何厂商资助开发的，[3]也不是被任何厂商策略性操控的：它们对所有人开放。行为主体就是技术的消费者，也是技术的直接或间接开发者。

假设行为主体i在时间t_i进入市场，在这一刻，他要在最新一代（即“最新版”）的技术A和技术B之间做出选择，然后使用这一版本的技术。[4]再假设行为主体有两种类型R和S，两种类型的行为主体的人数相同。类型独立于选择的时间，但两种类型的偏好不同，这也许是因为他们对自己所选择的技术的用法不同。每个行为主体所选择的A或B的版本，在他进行选择时是固定的或“冻结”在设计当中的，因而他的支付仅受到他所选中的技术以往被采用的历史的影响。在下文中，我将研究“包含预期”的情形，即支付同时还会受技术在将来被采用情况的影响。

当然，并不是所有的技术，都能够因为被采用而实现收益递增。在有的时候，投入的要素价格被抬高，就可能会出现技术随着被采用而递减的现象。

举例来说，水力发电可能会变得更加昂贵，因为随着水坝越建越多，坝址会变得更加稀缺和不适合。还有一些技术不受采用情况的影响，也就是说，它们的收益是恒定不变的。在我的模型中，所有这些情形都被包括进来了：假设任何一个行为主体选择技术A或技术B实现的收益，即那个版本的技术可以给他带来的净现值，取决于他进行选择那一刻以前的采用者的数量n_A和n_B（如表4-1所示）[5]，即到底是收益递增的还是收益递减的，抑或是收益不变的都随r和s而定，它们或者同时为正或者同时为负，又或者同时为0。我还假设，$a_R > b_R$且$a_S < b_S$，这也就是说，R型行为主体天生偏好技术A，S型行为主体天生偏好技术B。

表4-1　给定以往的采用情况时选择A或B的收益

	技术 A	技术 B
R 型行为主体	$a_R + rn_A$	$b_R + rn_B$
S 型行为主体	$a_s + sn_A$	$b_s + sn_B$

不确定性（**Uncertainty**）　指经济行为主体对于未来的经济状况，尤其是收益与损失的分布范围，以及状态不能确切知晓。

为了完成这个模型，我还必须仔细地定义“偶然”和“历史小事件”的含义。如果我们对可能影响技术选择的所有事件和所有情况，如政治利益、技术开发者先前的经验、订立合约的时机、关键性会议上的决定等都拥有无限详尽的先验知识，那么每种技术的结果或它在采用者市场中的份额，都可以提前确定。这当然不是事实。相反，我们可以得出结论，我们所拥有的有限的辨别能力，或者更确切地说，我们作为隐含的观察者的有限辨别能力，很可能导致结果的不确定性。因此，我将“历史小事件”定义为，那些超出了观察者的事前知识的事件或条件，即超出了观察者的“模型”，或他对情境的抽象“分辨能力”的那些事件或条件。

回到我们的模型。我们假设有这样一个观察者，他拥有关于所有条件和收益函数的完全知识，但是决定行为主体的进入时间和选择$\{t_i\}$的事件集除外。观察者“看到”的选择次序，将是R类型和S类型的行为主体组成的二元序列。在技术采用的“时间线”上，R或S出现在第n位的可能性相同，即概率均为1/2。

现在，我们有了一个简单的新古典配置模型。在这个模型中，两种类型的行为主体要在技术A和技术B之间做出选择，每个行为主体在轮到他时选择自己最喜欢的技术。供给函数或收益函数是已知的，需求函数也是已知的。每个行为主体都只需要一个单位，没有任何弹性。只有一个很小的因素是不确定的，那就是决定行为主体做出序列选择的那些历史事件的集合。我们感兴趣的是：第一，在收益递增、收益不变、收益递减这些不同的情况下，两种技术被采用的结果，即各自的市场份额如何；第二，当引入一些小事件，导致选择的顺序出现波动时，会不会对

各种技术的市场份额产生影响。

我们还需要进一步界定其他一些性质。对于一个过程，如果内嵌的小的不确定性可以被“平均化掉”，从而使得观察者可以获得足够的信息，准确地预先确定长期市场份额，那么我们就说这个过程是可预测的。如果对一种技术发放补贴或征税，总能影响未来的市场选择，那么我们就说这个过程是灵活的，而不是锁定的。如果历史事件的不同序列导致相同的市场结果的概率为1，那么就称这个过程是遍历的，而不是路径依赖的。在我们这个配置问题中，行为主体的选择确定了一条“路径”或技术A和技术B的各种版本被采用或被“开发”的序列，早期的采用者可能会将该过程引导到一个适合他们的开发路径上，但是该路径可能会被后来的采用者引为憾事。因此，和其他按时序进行选择的决策问题一样，在这里我将引入“不遗憾”标准，并规定在任何时候，如果技术的同等强度的开发或同样的采用率不会导致更好的收益，那么就说这个过程是路径有效（率）的。[6]

三种收益体制的配置

在研究我们这个由R型行为主体和S型行为主体组成模型的选择结果之前，我们不妨先来讨论一个简单的收益递增例子的动力学机制，即如果行为主体只有一种类型，又会怎么样呢？这种讨论可以给我们带来很大的启发。在这种情况下，选择的先后次序不再是一个问题，所有行为主体都是一样的。同时，未知事件也不会导致什么不同，所以遍历性也不是一个问题。假设第一个行为主体选择了他更喜欢的技术A，这个选择增大了采用技术A的收益；那么，下一个行为主体就更有理由选择技术A了，于是他也会选择技术A。这个过程将继续下去，每个行为主体每次都会选择技术A，技术B根本没机会“启动”。最终结果是，技术A“独占市场”，技术B则被排除在外。如果收益以相同的速率增加，那么这个结果就是可预测的，而且路径也是有效的。不过要注意的是，如果收益是以不同的速率增加的，那么采用过程就很容易变成路径无效的，如表4-2所示。

表4-2　同质行为主体的采用支付

以往采用的数量	0	10	20	30	40	50	60	70	80	90	100
技术*A*	10	11	12	13	14	15	16	17	18	19	20
技术*B*	4	7	10	13	16	19	22	25	28	31	34

由这个例子可知，在采用过程中进行了30次选择之后（前30次全都选择技术*A*），同样概率采用技术*B*会得到更高的收益。但是，如果采用过程已经“走了足够远”，那么即使给技术*B*某个给定的补贴*g*，也不可能再弥补该处技术*A*的收益与起点处技术*B*的收益之间的差距。因此，灵活性在这里不复存在，市场也变得越来越“锁定”于更劣的选择。

现在，再让我们回到最感兴趣的那种情况，即两种类型的行为主体的未知选择序列允许我们将“历史小事件”的概念包括进去讨论。我们先从收益不变的情况开始探析。用n_A（n）和n_B（n）分别表示当选择进行了n次时，技术*A*和技术*B*被选择的次数。这样一来，我们就可以用当选择总共进行了n次时，在第n阶段技术*A*所占的市场份额x_n来描述这个过程。为了方便起见，我们将采用数之差n_A（n）$-n_B$（n）记为d_n。于是，技术*A*的市场份额可以表示为：

$$x_n = 0.5 + d_n / 2n \quad (1)$$

不难看出，只要通过变量d_n和n（前者为采用数之差，后者为总和），我们就可以全面地描述技术*A*和技术*B*被采用的动力学了。在这种收益不变的情况下，*R*型行为主体总是选择技术*A*，*S*型行为主体总是选择技术*B*，而不管任何一种技术的采用者的数量是多少。因此，技术*A*和技术*B*采用者的累积数，是直接由*R*型行为主体和*S*型行为主体“排队”做出选择的顺序决定的。这就是说，如果排在下一位的是*R*型行为主体，那么n_A（n）就增加一个单位；如果排在下一位的是*S*型行为主体，那么n_B（n）就增加一个单位；同时两种技术被采用的数量的差d_n，也相应地向上移动一个单位或向下移动一个单位。对于某个观察者来说，选择顺序是随机的，而行为主体的类型是等概率的。因此，对于这个观察者来说，状态d_n所“执行”的似乎是一个简单的抛硬币式的随机游走：每一步都有相等的概率0.5。

而在收益递增的情况下，就没有这么简单了。这时出现了一种新的*R*型行为主体，他们尽管天然偏好技术*A*，但是如果偶然采用技术*B*，就可以在采用人数及收益上远远抛开技术*A*，那么他们就会“改换门庭”。这也就是说，新的*R*型行为主体将会使自己的偏好“切换”为更喜欢技术

B，即：

$$d_n = n_A(n) - n_B(n) < \Delta_R = \frac{(b_R - a_R)}{r} \qquad (2)$$

类似地，新的S型行为主体也会切换为偏好技术A，如果技术A的采用者的数量领先技术B的采用者足够多的话，即：

$$d_n = n_A(n) - n_B(n) > \Delta_S = \frac{(b_S - a_s)}{s} \qquad (3)$$

现在，选择区域是在d_n,n平面中（如图4-1所示），它们之间的屏障由式（2）和式（3）给出。一旦进入了某一个外部区域，这两种类型的行为主体就会都选择相同的技术，结果是该技术的领先优势进一步增大。因此，在d_n,n平面中，式（2）和式（3）描述了"吸引住"该过程的两条屏障。一旦d_n随机移动到了某条屏障，该过程就不再同时涉及两种技术，它会"锁定"到某一种技术上。因此，在收益递增的情况下，技术采用过程变成了一个有吸收屏障的随机游走。而在收益递减的情况下，从一名观察者的角度来看，上述分配过程是一个有反射屏障的随机游走，反射屏障可由两个类似于式（2）和式（3）的方程给出。相关的证明，我留给读者完成。

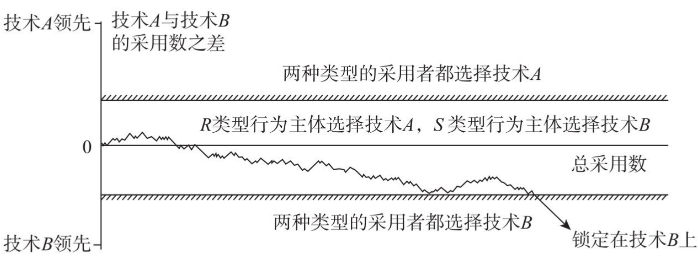


图4-1　收益递增情况下的技术采用

三种收益体制的特性

现在，我们可以运用随机游走的基本理论，来推导出上述选择过程在不同线性收益形式下的性质。为了方便读者参考，我把相关结果总结在了表4-3中。

为了证明这些性质，我们需要先考察技术的长期市场份额。在收益不变的情况下，市场是两种技术共享的。在这种情况下，随机游走的范围是自由的，但是我们从随机游走理论可知，d_n的标准偏差随着n增加。因此，式（1）中的$d_n/2n$项消失，同时x_n则趋向于0.5，从而使得市场分

割为50对50。在收益递减的情况下，市场份额也是共享的。

表4-3　三种收益体制下的性质

	可预测性	灵活性	遍历性	路径有效性
收益不变	是	不是	是	是
收益递减	是	是	是	是
收益递增	不是	不是	不是	不是

技术采用数之差d_n被限制在两个有限常数之间，因此当n趋向于无穷大时，$d_n/2n$趋向于0，而且x_n必定达到0.5。在这里，50对50的市场分割，就源于这些以相同的速率下降的结果。相反，在收益递增-吸收屏障的情况下，技术A的采用份额（市场份额）最终必将变为0或1。这是因为在一个吸收随机游走中，d_n最终必定会以概率1跨越屏障。因此，这两种技术不可能无限期共存：一种技术必定会排除另一种技术。

由此可见，可预测性在收益不变或收益递减的情况下是有保证的。在这两种情况下，市场份额将稳定在50对50的预测是正确的，其概率为1。然而，在收益递增的情况下，要想保证准确性，观察者必须预测技术A的最终份额为0或100%。但是任何一个预测都有一半的概率出错，可预测性因而丧失了。需要注意的是，尽管从理论上说，观察者可以预测某种技术占据整个市场的概率为$s（a_R-b_R）/［s（a_R-b_R）+r（b_s-a_s）］$，但是他不可能准确地预测实际的市场份额，无论他对供给和需求条件的知识有多少。

在收益不变的情况下，最多只有部分的灵活性。针对收益的政策调整在所有时候都可以影响选择，但前提是调整幅度要足够大，即足以填补对不同技术的偏好之间的差距。而在其他两种收益体制下，这类政策调整对应于一个或两个屏障的移动。在收益递减情况下，一个调整g总是会影响未来的选择，如果不是对市场份额有影响，就是对绝对数量有影响，因为反映屏障在未来会继续影响该过程的概率为1。因此，收益递减是有灵活性的。然而，在收益递增情况下，一旦这个过程被吸收到了技术A和技术B中，为了将屏障转移到足以影响选择所需的补贴或税收调整，也就是系统被“锁定”的程度的精确指数将无限增加。此时灵活性无法再保持下去。

在收益不变和收益递减的情况下，遍历性很容易得到证明。在收益不变时，只有在某种非常特别的“排队方式”下，如出现的S型行为主体永

远是R型行为主体的两倍，才能以0的相关概率偏离50对50的市场份额。而在收益递减时，任何历史事件序列，即行为主体的任何排队方式，都必定会将该过程保持在反射屏障之间，并将市场推向50对50的市场份额。在这两种情况下，行为主体都忘记了他们的“历史小事件”。但是，在收益递增这种收益体制下，情况却完全不同。行为主体形成的一部分序列导致市场结果“青睐”技术A，而剩余部分则导致市场结果“青睐”技术B。某些可能导致市场共享的、非常特别的排列，如S后紧跟着R，后面又紧跟着S、R，如此无限循环，其概率或测度为0。因此，确定｛t_i｝的小事件决定了市场份额的路径，这个过程是非遍历的或路径依赖的。也就是说，它是由它的历史小事件决定的。

在收益不变和收益递减的情况下，路径有效性很容易证明。在收益不变的情况下，以前的采用不会影响收益。每个类型的行为主体都选择自己偏好的技术，滞后技术没能得到进一步发展或进一步采用，也不意味着放弃一定的收益。在收益递减的情况下，如果一个行为主体选择了领先的技术，那么他必定更偏好可用的滞后技术。但是根据定义，进一步采用滞后技术又会降低他的支付。因此，选择导致技术采用过程沿着一条较劣的发展道路走下去，这个可能性不会出现。与这些情况相反，在收益递增的情况下，则完全可能导致较劣的发展道路。假设市场锁定在了技术A上，R型行为主体不会有什么损失。但是对于S型行为主体来说，如果他们偏好的技术B以往得到了平等的开发并且可供他们选择，那么每个S型行为主体就可以获得（b_S–a_S）。因此，这里至少有一种类型的行为主体会觉得“遗憾”。如果两种技术以不同的速度得到改进，那么无效率问题可能会进一步加剧。如果在早期选择了当初有吸引力但后来改进缓慢的技术，就可能将整个市场锁定在这个较劣的选择上。从长远来看，被排除的那种技术如果得到了同等的开发，那么两种类型的行为主体的境况都会得到改善。

模型的扩展

对于上面这个基本模型，我们不难在各个方向上进行扩展。同样的定性结果适用于M种技术竞争的情形，也适用于不同类型的行为主体所占比例不相同的情形。在这里，随机游走会发生“漂移”。而且，如果不同技术在不同的时间“到达”市场，那么动态过程仍然与以前一样，不过该过程将从n_A或n_B不为0的初始状态开始。因此，在现实世界中，早期启动的某种技术可能已经被锁定了，从而使得新的、可能更好的技术无法

在市场上获得一个立足点。

在行为主体的数量有限，且不会无限增多的情况下，吸收或反射及依赖于这两种性质的其他性质仍然成立，只要行为主体的数量与切换屏障之间的“间隙”相比足够大。

那么，对于由厂商资助开发的技术，策略性行动会不会改变上面描述的这些结果？对于这个问题，我们还没有完整的答案。

汉森（Hanson）在一个基于上述模型的模型中表明，在存在策略性行为的情况下，对技术的市场排斥也会发生：厂商会采取渗透定价法，即在早期承受一些损失来换取日后潜在的垄断利润，最终除了一个企业之外，所有其他企业都退出的概率为1。但是，在贴现率很高的情况下，企业最感兴趣的可能是当前的销售额，而不是让竞争对手关门大吉，这样市场分享现象就会重新出现。[7]

最有趣的一个扩展也许是理性预期的引入。在这种情况下，行为主体的收益受未来行为主体的选择的影响。这种情况的一个典型例子是标准（这种技术）的形成。在这个问题上，至关重要的是，未来的用户会不会选择我现在所选择的。卡茨（Katz）和夏皮罗（Shapiro）证明，在一个存在策略性互动的两阶段决策模型中，行为主体对这些未来选择的预期会使市场变得不稳定。我们可以将卡茨和夏皮罗研究结果推广到我们的随机动力学模型中。假设行为主体所形成的预期表现为他们对自己所身处的随机过程类型的信念。当产生于这些信念的实际的随机过程类型，与他们所认为的随机过程类型相同时，我们就可以得到一个理性预期（即得到了满足的）均衡过程。在收益递增的情况下，理性预期也会导致吸收随机游走，但是在这个随机游走中，对锁定的预期加速了锁定，从而缩窄了吸收屏障间距，并恶化了市场的根本不稳定性。

一般框架：考虑随机性小事件

如果有一个一般性的分析框架，可以在比上述基本模型更一般的前提假设和收益体制下，分析对技术的序贯选择问题，那无疑是非常有益的。特别是，如果我们能够确定在什么情况下，“为被采用而竞争”的技术市场最终必定会被某项单一技术所支配，那将是非常有用的。

在设计一个一般框架时，保留以下两个性质可能非常重要：（i）在可供选择的技术之间进行的选择，可能会受到选择时每种技术被采用的

数量的影响；（ii）“模型外”的小事件可能影响技术采用，因而必须允许随机性的存在。因此，技术采用的市场份额，可能不是直接决定了下一种被选中的技术，而是决定了每种技术被选中的概率。

接下来，我们考虑一个动力学系统。在这个动力学系统中，有K种技术，在每一次进行技术采用决策时，都有一种技术被采用，其概率分别为p_1（x）、p_2（x）、p_3（x）……p_K（x）。在这里，概率向量p是向量x的函数；而向量x则为到目前为止的采用总额中，技术1、技术2……技术K各自所占的采用份额（即比例）。初始比例向量给定为x_0。在这里，我把p（x）称为采用函数。

现在，我们可以研究这个动力学系统中，各种技术的长期市场份额或采用比例会是怎么样的。作为例子，考虑如图4-2所示的两个不同的采用函数（在这里，$K = 2$）。根据图4-2，我们可以推测，在采用过程中，当技术A的采用概率高于其市场份额时，技术A的比例倾向于上升；而当技术A的采用概率低于其市场份额时，技术A的比例则倾向于下降。如果采用比例或份额在总的采用数量增加时不再变动，那么我们就可以推测，它们在采用函数的某个不动点上“适应”下来了。

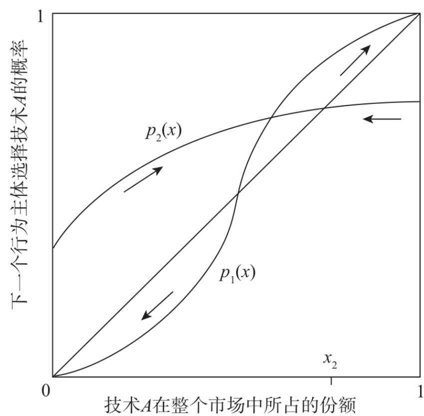


图4-2　采用函数示例

在1983年，阿瑟、埃尔莫利耶夫和卡尼奥弗斯基（Kaniovski）证明，在某些技术性条件下，上述猜想是正确的。[8]这种类型的随机过程，会以概率1收敛到从比例（采用份额）到采用概率的映射的某个不动点。当然，并不是所有的不动点都有这种“资格”。只有“吸引的”或稳定的不动点，即预期的运动过程所导向的点，才能够以长期结果的形式出现。而且，在采用函数随时间n变化，但趋近于限制函数p的情况下，该过程将收敛到p的某个吸引固定点。

因此，在图4-2中，可能的长期份额是0和1（对应函数p_1为x_1，对应函数p_2为x_2）。当然，在存在多个不动点的情况下，到底选择哪一个取决于该过程所采取的路径。这就是说，取决于随着过程展开而发生的随机事件的累积。

现在，我们有了一个通用的框架。根据这个框架，我们可以立即得到两个很有用的定理：路径依赖定理和单个技术支配定理（single-technology dominance）。

> 定理1：一个采用过程，当且仅当其采用函数p具有多个稳定的不动点时，是非遍历的和不可预测的。
>
> 定理2：一个采用过程，当且仅当其采用函数p只有在x是单位向量时才具有稳定的固定点时，以概率1收敛为单个技术支配。

这两个定理是前面阐述的基本理论的简单推论。因此，在两种技术竞争的情况下，只要在采用份额中至少存在一个不稳定的“分水岭”，当然必须存在多个不动点，那么该采用过程就是路径依赖的。在这个分水岭之上，技术的采用是自我强化型的，即倾向于增加自身的份额；在这个分水岭之下，技术的采用则是自我否定型的，即倾向于减少自身的份额。因此，技术在被采用后仅仅获得了一定优势是不够的。在特定市场份额上，这种优势还必须是能够自我强化的。

采用者类型连续时的非线性收益递增

作为一个例子，考虑上面的基本模型有一个更一般的版本，即采用者类型不只有两种而是连续的，他们在K种技术之间进行选择，同时他们支付的增加也可能是非线性的。假设如果先前的n_j个采用者以往选择了技术j，则下一个行为主体对采用技术j的支付为Π_j（n）$=a_j + r$（n_j），其中a表示行为主体对技术j的“天然偏好”，r是一个单调递增的改进函数，代表伴随着以前的采用者而来的技术改进。每个采用者对于这K种可供选择的技术，都有一个天然偏好向量$a=$（a_1,a_2，a_3……a_K），我们可以把这个行为主体想象为点a（它的支撑集是有界的）在正象限上的一个连续分布。由此我们假设，在每一次发生选择时，每个行为主体都是从这个概率分布中随机抽取出来的。单个技术j占据优势就相当于，支付Π的分布在采用过程的驱动下，以正的概率向一个能够使得

对于所有$i \neq j$且Π_j超过Π_i的点移动。

根据前述阿瑟-埃尔莫利耶夫-卡尼奥弗斯基定理，我们可以得到：

定理3：如果随着n_j的增加，改进函数r至少以速度ε增加，那么采用过程收敛为单个技术支配的概率为1。

证明在这种情况下，采用函数随总采纳数n的变化而变化。然而，我们不需要推导出它的确切形式。不难证明，随着n的增大：

（i）在采用份额的任一单位向量的邻域中的任何一个点上，无限的收益递增会导致相应的技术支配所有的选择，因此单位向量份额是稳定的不动点。（ii）等份额点也是一个不动点，但是不稳定。（iii）此外再没有其他点是不动点。因此，根据一般定理，由于极限采用函数仅在单位向量处有稳定的不动点，所以该过程以概率1收敛到这些不动点的某一个。而这也就保证了单个技术在长期中的支配地位。证毕。

然而，如果改进函数r是有界的，那么当学习效应耗尽时，单个技术的支配就不再是不可避免的了。这是因为，采用者类型的某些序列，可以使两种技术或更多技术同时得到或多或少的改进。然后，这些技术会一起达到r的上限，这就使得这些技术中没有一种能够占主导地位，市场将从那时起被它们所分享。与此不同，在其他采用者序列下，一种技术可能以足够快的速度达到上限，从而将其他技术“关在门外”。在有界的情况下，某些事件的历史动态地导致多种技术分享市场，其他事件的历史则会导致单个技术占据支配地位。总的来说，收益递增如果有界的话，通常不足以保证单个技术的最终垄断。

收益体制与路径锁定

实际经济在多大程度上被锁定在较劣的技术路径上？现在我们仍然不清楚。当然，相关的例子并不难找：某种早期技术占据了市场之后，导致后来出现的更优的替代技术无法在市场上获得立足之地。[9]关于导致锁定的历史事件，现在已经有了两个重要的安全研究：一是关于QWERTY布局的打字机键盘，二是交流电技术。在这两个案例中，收益递增都主要源于协调外部性。

还有其他一些案例也很有研究价值，它们反映了通过学习过程实现的路径锁定。这些案例包括，20世纪50年代和60年代的核反应堆技术的竞争，以及19世纪90年代的美国蒸汽动力车与汽油动力车的竞争。美国的核电站几乎清一色地采用了轻水反应堆技术，而这种核反应堆技术本来的目的是用来推进于1954年下水的第一只核潜艇“鹦鹉螺号”的，它的核反应堆高度紧凑。在美国核电技术的发展史上，一系列事件，包括海军的早期潜艇建筑合同、政治上的权宜之计、欧洲原子能共同体计划，以及关键人物的行动等，都有利于轻水反应堆技术。早期获得的学习和建造经验，似乎已经锁定了轻水反应堆技术的主导地位，关闭了其他反应堆技术发展的通道。然而，许多工程学文献都一致认为，如果能够得到平等的发展机会，那么气冷堆技术这种更加优越的技术是有可能发展起来的。又如，在汽油动力汽车与蒸汽动力汽车竞争的例子中，两种不同类型的开发者根据自己以前的经验，或偏好蒸汽动力，或偏好汽油动力，它们在不同时间进入了这个行业，并给出每种技术的最佳可用“版本”。在一开始时，汽油动力这个选择被认为是不太有希望的，因为汽油很容易爆炸，又显得比较脏，同时当时汽油发动机噪声也很大，而且需要复杂的新部件。[10]但是在美国，一系列看似琐碎、无关紧要的小事件，在世纪之交把几个主要开发者推向了汽油动力。然后到了1920年，汽油动力技术就把蒸汽动力技术排挤出了市场。尽管这两种技术孰优孰劣，在专业工程师之间仍然存在争议。

本章上面的论述表明，对不同收益体制，应该给出不同的经济史解释。在收益不变和收益递减的情况下，市场的进化只反映先验禀赋、偏好和转型的可能性，而且小事件不能动摇其结果。这似乎很能令我们放心，但是历史也就下降到了“只能充当承运人”的地位，只要将各种“不可避免的东西”交付出去就完事了。相反，在收益递增的情况下，各种各样的结果都是可能的。不确定的情况会因正反馈而被放大，从而“提示”系统进入实际“选定的”结果。这样一来，历史的小事件就变得很重要了。[11]因此，当我们观察到一种技术或一种经济结果压倒了它的竞争对手时，对那种执着于通过寻找赢家的“先天优越性”来解释它被采用的原因的做法，我们应该保持警惕。

通常的政策，即让更优越的技术通过在市场中占据支配地位这种结果来证明自己的优越性，在收益不变和收益递减的情况下是适当的。但是在收益递增的情况下，自由放任不一定能保证“更优”技术在长期意义上最终会生存下来。对于无资助的、通用的收益递增技术，应该根据市场细分的性质采取更有效的政策。在我们的模型中，早期采用者通过合

理选择只适合于他们自己的技术，对后来者产生了外部性。

在这种情况下，就缺乏一个行为主体间的市场，这个市场能够诱使人们去探索有前途的但更昂贵的“不成熟技术”，这样后来的采用者则可以轻松地受益。[12]在这种情况下，标准的补救措施是让早期开发者拥有获得一定补偿的权利，即专利制度，补偿则由后来的采用者支付，只有在早期开发者能够独占日后获益的范围内才可能是有效的。一种替代方案是，由某个政府权力机构根据某个有希望但不太受欢迎的技术路径，去对采用和开发进行担保。但是，如果最终的技术收益很难确定，如美国的战略防御计划，那么政府当局必然会面临一个经典的“多臂老虎机”难题：应该选择在哪些技术上“下注”。潜在的优越技术的早期失败结果，即低的“累积奖金”，可能导致政府当局“完全合理”地放弃它，转而采用其他可能的技术。这样一来，“被锁定在一个令人遗憾的发展过程上”这个根本问题依然存在。

动态收益递增会导致糟糕的“锁定”

本章考察了从多个“候选的均衡”中，“选出”一个均衡的动态过程。在这个过程中，各种经济力量与随机的“历史事件”相互作用，超越了通常的对收益递增问题的静态分析。在此我们阐明了，动态的收益递增是如何导致经济被逐渐地锁定到一个结果上，而这个结果不一定优于其他可能结果、不容易改变，并且不能完全被预测。

在收益递增的情况下，经济对象（在这种情况下是技术）之间的竞争呈现出一种进化特征，其机制类似于遗传学中的“创始者效应”机制。[13]在这里，“历史”变得重要了。在某种程度上，一个经济体中的技术发展，取决于观察者的模型无法“分辨”的小事件，因此我们很可能没有能力有任何把握去预测技术的市场份额。这就意味着，对于经济未来的可预测性，不但存在着实践上的限度，还存在着理论上的限度。

COMPLEXITY

AND THE

ECONOMY

05

经济中的过程与涌现

“复杂性视角”下的经济系统

复杂经济的显著特征告诉我们：传统经济学所使用的数学方法将面临极大挑战。把复杂性科学引入经济领域，将带来三个重大影响。所谓的“共同知识”，只能是通过具体的互动过程来获得经验，却无法用演绎推理方法获得。由行为主体之间重复互动模式所定义的网络结构，正成为新的研究热点。

本章由我和史蒂文·杜尔劳夫、戴维·莱恩合写于1997年，是圣塔菲研究所经济学研究项目的总结报告。事实上，这篇文章几乎就是一篇宣言。我的两个合作者史蒂文·杜尔劳夫和戴维·莱恩也对本文做出了很大的贡献。多年来，他们两人对复杂性和经济学有深入的思考。

我们认为，经济活动的行为主体是异质性的，他们并行地采取行动，以便应对他们共同创造出来的各种“聚合状态”。而在这样做的过程

中，他们需要不断地适应一个永恒变化的世界。在这个世界中，没有全局性的“控制者”；在这个世界中，人类的认知、层级结构和互动都是重要的。我们强烈指出，经济呈现出了复杂系统的所有标志性特征。这篇文章最初是《经济可看作是进化的复杂系统》第二卷的导论。通过这篇文章，就可以对当时的复杂性科学研究有一个很好的了解。

* * *

1987年9月，20位学者齐聚圣塔菲研究所，讨论有关“经济可看作是进化的复杂系统”的问题。在他们当中，有10位是理论经济学家，受肯尼斯·阿罗邀请而来；另10位则是物理学家、生物学家和计算机科学家，受菲利普·安德森邀请而来。在当时，自然科学中已经涌现出了一系列新思想，它们松散地集合在了“复杂性科学”这个概念之下。举办这次会议的动机就是，探讨这些新思想，能不能激发出新的思考经济问题的方法。会议整整进行了10天，经济学家和自然科学家“轮番上阵”，报告他们各自的“世界”和“方法”。物理学家了解了一般均衡分析和非合作博弈理论，而经济学家则试图理解自旋玻璃模型、布尔网络和遗传算法。

这次会议产生了两大成果。第一个成果是一本文集《经济可看作是进化的复杂系统》第一卷，由阿罗、安德森和戴维·皮内斯主编。另一个成果是1988年在圣塔菲研究所启动的经济学研究计划，它也是圣塔菲研究所的第一个常驻研究计划。这个研究计划的任务是，从复杂科学的角度增进对经济现象的理解，具体内容则包括两部分：一是发展新理论，二是探索用来建模和实证分析的新工具。自1988年以来，这个研究计划已经邀请了很多杰出的学者来圣塔菲研究所参与研究工作。它还资助了一系列研究项目，每年都举办多次研讨会，并完成了几十篇工作论文。此外，自1994年以来，它还为经济学研究生举办一年一度的暑期学校。

1996年8月，圣塔菲研究所经济研究计划举办了一个研讨会，《经济可看作是进化的复杂系统》第二卷（以下简称“第二卷”）就是会议的重要成果之一。这个研讨会的目的是要回答这样一个问题：在过去10年里，复杂性方法对经济学有什么贡献？与1987年那次会议后出版的文集不同，第二卷中所有文章几乎都是讨论经济问题的，而且大多数作者都是经过正规经济学训练的经济学家。此外，虽然有相当一部分文章是圣塔菲研究所构思或实施研究工作的成果，但是也有不少文章的作者在此之前与圣塔菲研究所没有任何联系。事实上，与复杂科学相关的研究，已经在世界各地的许多不同的研究机构和大学中展开。

但是，在经济学中，“复杂性科学视角”或“复杂性观点”（complexity perspective）到底是指什么？这并不是一个容易回答的问题。“复杂经济学观点”的意义仍然在构建中，现在并未完成。第二卷的目标就在于促进这个构建过程。当然，第二卷各篇文章的作者仍然没有对经济学中的复杂性意义和重要性达成一个单一的、内在一致的认识。但是，我们确实可以找到一组相互接近、相互关联的主题，它们共同构建了经济学中的复杂性视角在当下此刻的含义。

其中有几个主题从20世纪80年代中期以来，就已经成了许多经济学家研究的重要课题，它们在前期出版的《经济可看作是进化的复杂系统》第一卷中也得到了很好的讨论。具体来说，非线性动力学在经济理论和数据分析中的应用，在1987年的会议上由米歇尔·博俊（Michele Boldrin）和威廉·布洛克（William Brock）报告；正反馈理论，以及与之相关的关于路径依赖和锁定的分析，则由布莱恩·阿瑟报告；1987年后，与这两个主题相关的研究，已经在圣塔菲研究所经济研究计划，以及其他研究机构蓬勃开展起来了。在1987年时，混沌是非线性动力学中人们最有兴趣的核心概念，不过后来它逐渐淡出了中心舞台。过去10年来，经济学家在识别金融时间序列中的非线性模式方面，已经取得了重大进展。他们提出了一系列模型，用来解释和分析这些模式，并在某种程度上预测这些模式得以呈现的序列。第二卷收录的布洛克的文章，对这方面的研究进展进行了综述。正反馈在莱恩的关于信息传播的模型、杜尔劳夫的关于不平等的模型和克鲁格曼的关于经济地理的模型中都发挥着核心作用，而且也是隐藏在诺思（North）所描述的经济发展现象和莱永胡武德（Leijonhufvud）所描述的恶性通货膨胀现象背后的“幽灵”。

回顾过去10年来这些领域的发展和圣塔菲研究所经济研究计划所产生的成果，我们认为一个连贯的经济学观点，它有时被称为“圣塔菲观点”，已经出现在经济学中了。在这里，我们将它称为“复杂性视角”，或称为“圣塔菲研究所视角”，偶尔也称之为“过程与涌现视角”。在我们细致地对它进行描述之前，我们要先描述一下作为标准的新古典经济学基础的两个基本概念。理论经济学家们在1987年那次会议上的大多数发言都属于这种经济学。我们可以将这两个基本概念，称为“均衡”和“动态系统”。从均衡观点出发，经济学家感兴趣的问题是，如何从个体最大化者的理性选择中，导出总体层面的某种“经济状态”，如一般均衡分析中的一组价格、博弈论中的一组策略及相关的支付，导出满足某些总体层面的一致性条件，如市场出清、纳什均衡，并考察这些总体层面的经济状

态的性质。而在“动态系统”方法中，先用一组变量来表示经济状态，然后用差分方程组或微分方程组描述这些变量如何随时间变化。运用“动态系统”分析时，经济学家感兴趣的问题是，考察映射在状态空间上的结果轨迹。但是关键在于，均衡方法未能描述经济状态随时间变化的机制，也没有描述均衡是如何形成的。[1]“动态系统”方法则通常不能同时容纳行为主体（个体）层面和系统（总体）层面的分析，无法填补两者之间的鸿沟，或者只能通过“代表性行为主体”这种工具来遮掩这个问题，而且也不能解释新的相关状态变量的出现，更不用说新的实体、新的模式和新的结构涌现了。[2]

复杂经济的6大特征

为了更加清晰地描述复杂性方法，下面我们先指出经济的6个特征。这些特征共同决定了，仅使用传统经济学使用的数学方法将面临极大的困难[3]：

分散的交互作用。经济中发生的几乎所有事情，都是由大量的行为主体的相互作用决定的，而这些行为主体是分散的且极可能是异质性的、并行采取行动的。任何一个特定的行为主体的行动，都依赖于他所预期的一定数量的其他行为主体的行动，还依赖于所有行为主体共同创造的聚合状态或总状态（aggregate state）。

没有全局性的控制者。经济中没有全局性的、可以控制行为主体之间的交互作用的实体。如果说有控制，那也是通过行为主体之间的竞争和协调机制实现的。经济行为是由法律制度、行为主体承担的角色，以及相互之间不断变化的联结来实现调和的。同样地，经济中也不存在万能竞争者，即可以利用经济中的一切机会的一个行为主体。

交叉分层组织。在经济中，存在着很多层次的组织和互动。任何一个给定层级的“单元”，通常都要成为构建下一个更高层级的单元的“构件”。经济的整个组织不仅是层次化的，而且在每个层次上都存在许多复杂缠结的交互作用，或者说联系或交流渠道。

连续适应。随着行为主体经验的不断积累，行为、行动、策略、产品等，全都不断地被修订、调整，或者说系统将不断地适应。

永恒的创新。随着新的市场、新的技术、新的行为模式、新的机构等不断涌现，新的“利基”不断地被创造出来。而且，填补一个利基的行

动本身也可能会创造新的利基。这导致的结果是持续的、永恒的创新。

非均衡的动力学。由于新的利基、新的潜力、新的可能性等不断地被创造出来，经济的运行会远离任何最优或全局均衡。改进总是可能的，并且确实经常发生。

具有这些性质的系统，通常被称为“自适应非线性网络”。这个术语是约翰·霍兰德创造的。在自然界和人类社会中，这种自适应非线性网络随处可见：神经系统、免疫系统、生态系统，以及经济系统。自适应非线性网络的一个基本要素是，它们不是简单地以刺激-反应的形式采取行动。恰恰相反，它们有预期。特别是，经济行为主体会形成预期。他们会建立经济模型，并利用这些模型来进行预测，然后在预测的基础上采取行动。注意，这些“预期模型”既不一定要明确，也不一定要连贯，甚至也不一定要相互一致。

由于经济的上述特征带来的困难，如果只运用经济学家习惯使用的数学工具，如线性、不动点、微分方程组等，就不能对自适应非线性网络有一个深刻的理解。相反，我们需要的是，将组合数学和种群层级的随机过程结合起来的新型数学方法，然后再加上计算机建模方法。这些数学方法和计算技术现在还处于起步阶段。它们都强调组织的各个不同层级上的结构发现和结构涌现的过程。

“复杂性视角”带来的三大影响

“作为一个自适应非线性网络的经济”这个概念，或者说“作为一个不断进化的复杂系统的经济”这个概念，对经济理论的基础、对提出和解决经济学理论问题的方式，必将产生深远的影响。接下来，我们就来阐述这些影响。

认知基础。新古典经济学理论有一个一元化的认知基础：经济行为主体是理性的最优化者。根据通常的解释，这就意味着，行为主体以概率的方式评估不确定性，然后根据以贝叶斯方式更新的新信息来修正他们的评估，并且选择能够使他们的期望效用最大化的行动方案。在这个统一的一元化的认知基础之上，行为主体通常被假设为拥有关于对方的共同知识，还会对他们所处的世界，当然也是他们共同创造的世界做出理性预期。相比之下，圣塔菲的观点无疑更加多元化。遵循现代认知理论，我们不再假设认知过程有一个单一的、主导的模式。相反，我们认为行为主体必须在认知上“构造”他们面临的问题，即他们必须理解他们

的问题、给他们的问题赋予“意义”，然后才能解决他们的问题。他们必须利用有限的认知资源来做到这一点。为了“有意义”、为了学习和适应，行为主体要运用各种各样的分布式认知过程。行为主体用来将自己关于世界的信息转换为行动的范畴，源于他们自己的经验，并且这些范畴，或者说认知“道具”不一定彼此一致也能产生有效的行动。因此，行为主体身处的是，一个他们必须“认知地”加以解释的世界，一个由于其他行为主体的存在和行动而复杂化，并且不断变化的世界。这也就是说，行为主体之所以通常不会进行标准经济学意义上的最优化，并不是因为它们受到“记忆能力有限”或“处理能力有限”的限制，而是因为最优行动这个概念本身通常无法明确地定义。据此，更进一步的结论也就水落石出了：说起如何指导人们在这个世界上的行动，新古典经济学中行为主体的演绎理性，最多只能占据边缘位置。也正因为如此，行为主体可能拥有的任何关于彼此的“共同知识”，都必定源于具体的、特异性的认知过程。而这种认知过程所处理的，则是通过具体的互动获得的经验。这就是说，我们不能简单地存在共同知识。

结构基础。在一般均衡分析中，行为主体不直接彼此互动，而要通过非人格化的市场。相比之下，在博弈论中，所有博弈参与人都与所有其他博弈参与人互动，而结果则用博弈的支付矩阵来给出。在这两种情况下，互动结构都是简单的且通常是极端化的：一个人与所有人，或者所有人与所有人。此外，行为主体自身的内部结构也被完全抽象化了。[4]相反，从复杂性的视角来看，结构是非常重要的。第一，基于网络的结构变得很重要。所有的经济行动都涉及行为主体之间的相互作用，因此经济功能不仅受行为主体之间重复互动的模式所定义的网络约束，同时也是在这个网络上实现的。这些网络结构的特征则是用网络连接的相对疏密程度来刻画的。第二，经济行动是通过涌现的社会角色和社会所支持的程序，也就是说，通过制度来完成结构化的。第三，经济实体具有递归结构，它们本身又是由其他实体构成的。不过，由此而形成的实体“层级”结构及其相关联的行动过程，却不一定是严格等级化的，因为作为组件的实体可以成为不止一个更高层级的实体的一部分，而且位于不同组织层级的实体之间也可以相互作用。因此，互惠因果关系会在组织的多个层次之间发挥作用。这就是说，尽管在给定的某个组织层次上的行动过程，有时可以被视为自主的，但是它们还是要受到其他层次的行动模式和实体结构的约束。而且，他们甚至有可能在更高和更低的层次上产生新的模式和实体。从“圣塔菲观点”的角度来看，组织的基本原则是，一个层次上的“单元”结合在一起，产生下一个更高层次上的“单

元”。[5]

过程与涌现。现在应该很清楚了，从我们上述的观点来看，简单地直接将经济事务视为多行为主体的最优化努力，并没有多大意义。我们的观点强调的是过程而不仅仅是结果。具体来说，它旨在探究新的“事物”如何在世界中生成。这包括：认知事物，如“内部模型”；物理事物，如“新技术”；社会事物，如新型的经济“单位”。很显然，如果我们假设一个“永恒新异”的世界，无论是瓦尔拉一般均衡、纳什均衡，还是“动态系统”理论中的均衡，那么结果就不可能对应于一个稳态均衡。在这样一个世界中，唯一可能有效的描述，只能是关于瞬变现象的描述，即关于过程和涌现结构的描述。那么，从这种“过程与涌现”的视角出发，我们对经济能够知道些什么？又应该怎样去理解经济呢？对经济中的过程与涌现的研究，现在已经形成了一个颇为壮观且仍在成长的“行业”，它通常被称为“基于行为主体的建模”。不过直到现在，什么才算一个基于行为主体的模型的解，仍然是一个聚讼纷纭的问题。第二卷所收录的许多文章，包括阿瑟和达利（Darley）、考夫曼、舒比克（Shubik）、林德伦（Lindgren）、科尔曼（Kollman）等人，以及基尔曼和特斯法齐的文章，都或明确或隐含地讨论了这个问题。总的来说，这些文章都试图在互动过程中寻找涌现结构，而在这种互动过程中，相互作用的经济实体通过各自的认知过程来预测未来。认知过程本身又会涉及发生于多层级结构中的互动。

当然，对一种经济学方法的描述，并不一定能构成一种研究纲领。要想围绕“过程与涌现观点”提出一个研究纲领，还必须完成两件事情。第一，必须确定具体的经济问题，这种方法可以对这类经济问题提供新的洞见。第二卷给出了这类经济问题的若干“候选者”：人工制品创新、贸易网络的进化、货币的产生、城市的起源和空间分布、资产定价、恶性通货膨胀，以及不同社区或国家之间的长期收入差异。第二，对这些问题建模所必需的认知基础和结构基础必须先构建起来，同时能够将基于前述基础的各种理论和可观察的现象联系起来的方法也必须先发展出来。虽然自1987年以来，我们已经取得了重大进展，但是这样一个研究纲领仍然远未形成。

三个新热点

第二卷收录的各篇文章，可以说都是对交互式世界中的过程与涌现

这个中心主题的探索。这些探索是平行进行的，它们为如何研究那种能够产生永恒的新异性的系统提供了很好的线索。当然，这些探索并没有形成一个连贯的整体。它们基本上是互补的关系，有时甚至可能是存在局部矛盾的。但是，要想描述“圣塔菲观点”的话，还有什么能够比强调分布式过程、涌现和自组织更加贴切的呢？在这种探索过程中，一些研究方向已经涌现出来了，我们下面就来总结一下。

认知。第二卷提出的核心认知问题都是解释问题。正如舒比克所指出的：“对数据的解释至关重要，关键不在于数据是什么，而是数据的意义。”那么，行为主体要怎样做，才能使他们的世界变成可理解的，从而使得信息“有意义”呢？阿瑟、霍兰德、勒巴伦、帕尔默和泰勒的文章，以及达利和考夫曼的文章也都强调了这一点。这两篇文章考察了一组行为主体如何采取行动的问题，他们的行动的影响取决于其他行为主体做什么。行为主体基于自己形成的有关其他行为主体将如何行动的预期，来决定如何行动。那么，这些预期来自哪里？这两篇文章都拒绝将共同知识或共同预期作为分析的起点。事实上，阿瑟等人在文章中指出，共同信念无法用演绎推理的方法得到。因为行为主体必须从一个想象中的未来得出他们的预期，而这个想象中的未来，正是其他行为主体预期的聚合结果或“总结果”。因此，预期是自我指涉的，这将导致演绎推理的不确定性。相反，这两篇文章都假设，每个行为主体都可以“访问”各种各样的“解释器”。这些解释器能够将这个世界的某些特定元素筛选出来、赋予它们意义，并能根据这些元素所传达的“信息”提出关于何种行动方为有效的建议。行为主体会把这些解释器的“表现”记录下来，舍弃那些提出“不好建议”的没有用的解释器，并对那些有用的解释器加以调整和改进。根据这种观点，经济行为产生于一个由解释器构成的不断进化的生态。在这个生态中，解释器以行为主体为中介相互作用，而行为主体则利用解释器去生成他们的预期。

在上述思想的基础上，阿瑟等人提出了一个资产定价理论。在他们的模型中，行为主体就是投资者，以“市场统计师”的身份采取行动。这些市场统计师不断地生成预期模型，即对市场上价格变动的解释，并通过交易来检验它们。如果不成功，他们就丢弃旧模型，用新模型取而代之。这样一来，市场的预期也就变成是内生的了。它们不断地改变、调整以适应它们共同创造的市场。因此，阿瑟等人的市场是一个拥有丰富多彩的“心理”的市场，在那里会涌现出投机泡沫、技术交易和持续的波动。而标准经济学文献所假设的同质理性预期则成了一种特殊情况，虽然它在理论上是可能的，但是在实践中却不太可能出现。布洛克引入了

这种方法的一个变体，允许行为主体在数量有限的预期模型之间切换。他的模型比阿瑟等人的模型简单，不过他得到了解析结果。布洛克的模型可以解释金融时间序列数据的许多特征事实，其中许多已经是学者们在过去10年中应用非线性工具进行分析时发现过的。

而在达利和考夫曼的模型中，行为主体分布在一个网格上，他们试图预测在网格上的“邻居”的行为。在生成预测时，这些行为主体使用的是一个自回归模型，而且他们每个人都可以单独地调整模型中参数的个数，以及他们用于估计模型参数的时间序列的长度。具体做法是，在每个周期，如果在前一期进行这种调整可以产生更好的预测的话，他们就可以按1的步长改变参数数量或历史长度。这样一来，就引入了一个共同进化的“解释动力学”，系统不可能进入一个有稳定的、精准协调的共同预期的状态。特别是，当系统接近于“稳定的理性预期状态”时，就会趋向于裂解为无序状态。根据这些结果，达利和考夫曼对具有无限的前瞻视野和无限的演绎能力的传统理性观念提出了挑战。

在他研究恶性通货膨胀的文章中，莱永胡武德提出了与达利和考夫曼相同的问题：在“视野无限长远的最优化”与“近视的适应”这两个极端之间，我们应该如何定位行为主体的认知能力？莱永胡武德认为，这个问题的答案是“情境依赖”的。他声称，在像恶性通货膨胀这样制度崩溃的情况下，行为主体的认知会转向“短时记忆／近视的适应模式”。因此，制度转换与认知模式转变之间的因果关系就成了“相互性”的。随着前瞻视野的萎缩，长期贷款（在他的模型中，长期意味着超过15天）的市场消失了。而且，当通货膨胀加剧后，会计核算的单位也会失去意义。因为不能以任何有意义的方式制定预算了，政府的行政分支无法继续在财政上对议会负责，地方政府也无法继续在财政上向中央政府负责。社会和经济的控制机制遭到了侵蚀。部长们失去了对他们的官僚机构的控制，股东也失去了对公司管理层的控制。

“解释器”，如清晰的预测模型和技术交易规则，在行为主体的认知中发挥着核心作用，这个思想与认知科学中更一般的一系列思想非常契合。对此，克拉克（Clark）在他的著作中进行了很好的总结。克拉克反对全部认知功能都集中在“头部”的观点。他认为，大量的解释性辅助工具，如自回归模型、计算机、语言，甚至如哈钦斯（Hutchins）所指出的导航工具及制度等，提供了一个“脚手架”或者说是“外部结构”，完成了大部分解释世界的任务。克拉克还进一步指出，即使是“头部认知”，其独特标志也是“快速的模式完成”，这与新古典主义经济学家孜孜不倦

所追求的演绎理性无关。在第二卷中，诺斯的文章也讨论了这个主题。他描述了以制度为“脚手架”来为解释提供支撑的若干方式。目的是解释对于经济主体来说，什么样的行为是可能的和适当的。

莱恩和麦克斯菲尔德在他们的文章中，从不同的角度探讨了解释问题。他们对所谓的“功能的归属”特别感兴趣。他们所称的“功能的归因”，就是指解释一种人工产品是“做什么的”。他们认为，新的功能归因出现在行为主体之间的、某些特定类型的互动情境中。在这种情境中，行为主体的解释可以不同。因此，认知有一个不可避免的社会维度。什么解释是可能的，取决于谁与谁互动以及为什么进行互动。他们还指出，新的功能归因是不可能在它们得以出现的特定生成关系之外预见到的。在行为主体-人工制品空间的结构快速变动的情境中，这种不可预见性对解释构成“理性”行动的是什么有深刻的影响。

上面提到的所有文章都强调了认知对经济理论的根本重要性。但是，一种相反的观点也可以从过程与涌现角度合理地加以辩护。根据这种观点，过度认知构成了源于方法论个人主义的另一个错误，方法论个人主义是标准经济理论的基石。单个行为主体如何决定做什么可能并不是太重要。发生的事情尽管被视为他们行动的结果，但很可能更多地依赖于他们行动的互动结构，即谁在互动、他根据哪些规则与谁互动。布卢姆在关于种群博弈的文章的引言中提到了这一点。正如他所指出的，他提供的模型所关注的，将不再是“个人层面的决策理论的细节，而是行为主体互动的动力学”。帕吉特（Padgett）也得出了类似的结论，尽管理由有所不同。帕吉特感兴趣的是，提出一个企业理论，作为变革性的“工作”的基点。他认为“工作”可以表示为，“一系列行动和反应组成的序列，这个序列可能会带来一些意图中的或非意图中的集体结果”。因此他指出，只要研究协调好的行动-反应序列的结构，就可以洞察经济活动的组织。他的“故事”中，根本没有引入“认知”。帕吉特文章的灵感来自于化学和生物学领域的一些最新研究成果，如艾根（Eigen）和舒斯特的研究、冯塔纳和布斯的研究，以及其他一些人的研究。它们堪称从复杂性视角出发进行研究的典范。

结构。绝大多数人类互动，当然也包括那些发生在“经济”环境中的人类互动，都有社会性的，这是人类互动首要的性质。与朋友交谈、向有常识的熟人征求建议、与同事一起工作、与邻居一起生活等，这种社会交往连续地、循环地进行，将行为主体“绑定”到了网络当中。[6]但是根据标准的经济学理论，行为主体做什么，只取决于他们自己的价值观

和可用的信息。而且标准理论通常忽略了价值和信息来自哪里这个问题。它将行为主体的价值观和信息视为外生的和自主的。在现实中，行为主体要相互学习，他们的价值观可能受他人的价值观和行为的影响。这种相互学习和相互影响的过程，是通过行为主体所嵌入的社交网络进行的。它们可能产生重要的经济后果。例如，杜尔劳夫在他的文章中提出的一个模型，意味着邻居之间的"价值关系"可能导致邻居之间的持续收入不平等。而在莱恩的模型中，网络中流动于行为主体之间的信息决定了两个相互竞争产品的市场份额。基尔曼的文章则回顾了一系列讨论社交网络如何导致一定的经济后果的模型。

伊奥尼迪斯（Ioannides）、基尔曼和特斯法齐考察了网络如何从最初随机的二元相互作用模式中涌现出来，以及这种涌现的网络会呈现出什么结构这个问题。伊奥尼迪斯还研究了基于受控随机场的数学模型。而在特斯法齐的基于行为主体的模型中，"行为主体"实际上是囚徒困境博弈中的不同策略。伊奥尼迪斯和特斯法齐所关注的，主要是与经济互动有关的网络形成和深化问题，特别是贸易网络。激发他们进行这些研究的思想，在一些社会学家，如贝克（Baker）的圈子内，很久以前就已经得到了认可，但在经济学界则很新颖，这个思想就是市场实际上是通过交易者组成的网络运行的。在市场中发生的事情可能反映了这种网络的结构，而网络的结构又可能取决于网络是如何涌现出来的。

局部性的相互作用可能导致大规模的空间结构的涌现。第二卷中有好几篇文章就是研究这个现象的。从这个角度来看，林格伦的文章特别有意思。像特斯法齐一样，他也运用了一个基于行为主体的模型，在这个模型中，行为主体要制定两人博弈中的策略。在林格伦的模型和特斯法齐的模型中，依据他们过去与其他行为主体博弈时所采用的策略的成功程度，行为主体随着时间的推移来调整自己的策略。不过，与特斯法齐的模型中的行为主体不同，它们随机相遇并能够自行决定是否展开互动，林格伦的模型中的行为主体只能与预先指定的交互网络中的邻居互动。利用这个模型，林格伦研究了行为主体空间上时空结构的涌现。这种结构是一种由一组策略构成的亚稳定生态：连续很多世代，这些策略能够抵御新的策略类型的"入侵"，或者说这种生态能够在自己的空间边界上与其他生态相互"竞争"。特别是，林格伦对网格网络中涌现出来的结构，即每个行为主体只与几个其他行为主体互动，与完全联接的网络中涌现出来的结构，即每个行为主体都与所有其他行为主体互动，作了详细的对比。结果发现，前者"导致若干策略之间的稳定共存，但如果没有这种网络结构，这些策略将会在竞争中被淘汰。这些时空结构呈现出

来的形式，可以是螺旋波、不规则波、时空混沌、冻结块状模式，以及各种不同的多种多样的几何构型”。虽然我们还不能毫无争议地说林格伦的模型是一个经济学模型，但是他对两种行为主体空间上的结构的比较，却无疑是非常有启发性的：一种行为主体空间中的社会网络是相对稀疏的，而在另一种行为主体空间中，所有行为主体间的互动都是有可能的。后者至少在原则上，可以与经济学中的一般均衡分析所特有的非人格化市场相比拟。帕吉特在他的文章中也进行了类似的比较，不过是在一个完全不同的情境下。

杜尔劳夫和克鲁格曼都研究了地理隔离是怎样出现问题的。在他们的模型中，行为主体可以改变位置，即改变他们在社会结构中的位置，而社会结构是由邻居关系定义的。在这些模型中，特别是在杜尔劳夫的模型中，存在着多种类型的行为主体。模型要回答的问题是：在什么情况下、通过什么机制，总体层面上的“邻居关系”或“社区”会涌现出来。每个社区都主要或者完全由一个类型的行为主体占据。因此，当前的网络结构，即行为主体的邻居，以及行为主体可以移动到“站点”处的邻居，调节着行为主体的选择，而行为主体的选择又会改变当前的网络结构。这样一来，随着时间的推移，从不断变化的局部网络结构中，总体层面上的相互隔离的社区就涌现出来了。

科尔曼、米勒和佩奇[18]也探讨了类似的主题。他们的文章的现实背景是，当存在多个司法管辖区时的政党竞争和民主制度问题。在他们的基于行为主体的模型中，行为主体可以在司法管辖区之间迁移。他们证明，当存在三个以上的司法管辖区时，“两党制”式的竞争优于民主的全民公决；相反的情形仅出现于只存在一个司法管辖区的时候，因此行为主体无法迁移。他们还发现，“两党制”式的竞争会比民主的全民公决导致更多行为主体进行迁移。

曼斯基（Manski）的文章则提醒我们，即使理论模型已经做得尽善尽美了，对真实现象的理解也同样重要。他区分了学者们对于人们经常观察到的经验事实的三种因果性解释：“属于同一群体的人倾向于采取类似的行动。”第一种解释是我们上面已经描述过的，即行为相似性可能通过网络交互效应产生。但是还有另外两种可能的解释。一种是环境，即这类行为可能取决于群体的某个外生特征，如社会经济构成。另一种是相关效应，即行为相似性可能源于群组成员共享某种相似的个体特征。曼斯基还证明，除了其他结果之外，使用流行的平均线性模型去分析数据和“观察均衡结果和参照群组构成”的研究者，是不能凭经验从上面这

些可供选择的解释中区分出内生性交互作用的。这里有一个“底线要求”，那就是非线性效应需要非线性的推断技术。

在诺斯、舒比克和莱永胡武德各自的文章中，关注的焦点转移到了另一种社会结构上，那就是制度。诺斯的文章的侧重点是研究制度和经济增长，舒比克则志在分析金融制度，而莱永胡武德则阐述了恶性通货膨胀的“现象学”。

所有这三位学者都同意，可以将制度定义为“游戏规则”，如果没有这些规则，经济行动将是不可想象的。不过，经过具体分析，他们至少在以下三个意义上使用“制度”一词：作为“规则”本身，如破产法；作为被授予了颁布规则的社会权力和政治权力的实体，如政府和法院；以及作为有社会合法性的建构，它们将规则“实例化”，而且经济行为主体就是通过它们来采取行动的，如法定货币、市场。无论在哪一种意义上，三位作者都认为，如果仅仅从纯粹的经济、纯粹的政治或纯粹的社会角度来进行分析，制度是不可能被充分理解的。经济学、政治学和社会学的作用，在制度形成的过程中是不可分割的。而且，制度也改变和决定了经济行为、政治行动和社会运动。诺斯还进一步坚持认为，制度拥有一个认知维度，它体现在维持着制度并决定了制度进化方向的、总体层面的“信念系统”上。

诺斯从功能主义的角度分析了制度如何涌现的问题：制度之所以出现，是“为了减少不确定性”，这就是说，制度能够使行为主体面对的世界变得更加可预测，从而为行为主体提供可识别的、能采取有效行动的机会。特别是，现代经济依赖于能够降低非人格化市场中交易成本的制度。

相比之下，舒比克则采取了一种不同的分析路径。他的分析从他的策略性市场博弈的概念开始。这种博弈是有“完全定义的过程模型”，它们指明了“可行结果集中所有点上可以采取的行动”。舒比克对使用法定货币进行贸易的经济，构建了一个策略市场博弈。他证明，指定可以采取行动的“完全规范”，就意味着某些类型的规则在逻辑上的必要性。舒比克把这种规则等同于制度。吉纳科普洛斯（Geanakoplos）在他的文章中，也提出了与舒比克类似的观点。他认为，金融制度就代表了承诺。那么，如果有人不能或不愿兑现承诺，又会发生什么？舒比克证明了其中一种制度，即破产法，在处理违约问题时的逻辑必要性。吉纳科普洛斯则引入了另一种制度：抵押品。他证明，在均衡中，作为一种制度的抵押品具有制度含义，即市场缺失。

最后，在他为第二卷撰写的结语中，菲利普·安德森从一个物理学家的角度，对费尔南·布罗代尔（Fernand Braudel）所强调的“长期社会经济变革的历史，带给我们一个最核心的教训”进行了评论。关于行为主体同质性的假设，在复杂系统中可能有非常强的欺骗性。社会变迁过程通常是由位于某些有“极端尾巴”的分布的最末端的那些人所驱动的。因此，从“圣塔菲观点”出发，一个有趣的理论问题就是：这些有极端尾巴的分布是怎么涌现出来的？为什么这种分布会如此无处不在，如此重要？

问题是什么？问题的解是什么？虽然第二卷收录的文章，有很多是讨论认知和结构的，但是它们并没有深入探讨：从上述新观点出发，什么可以构成一个问题？什么东西可以算作问题的解？也许这是因为，对生成和评估“解”的各种方法的讨论现在还为时尚早。我们还在讨论，我们应当致力于对什么东西求“解”。第二卷中有少数几篇文章完全采用了文字表述，不过大多数文章都提出了数学模型，包括基于统计力学的模型、策略性市场博弈模型、随机图模型、种群博弈模型、随机动力学模型，以及基于行为主体的计算模型，等等。需要指出的是，在某些情况下，有些文章作者所使用的数学模型忽略了一些重要的问题。例如，均衡计算能够以什么方式提供关于涌现问题的洞见？这是一个相当棘手的问题，但是没有任何一篇文章讨论了这个问题，即便是那些用于计算均衡的模型，并声称提出了关于涌现问题新见解的文章也是这样。布卢姆在他对种群博弈模型的讨论中，提出了两个与此相关的问题：在他的理论中，渐近均衡选择定理能不能保证均衡选择“发生得足够快”，从而使它在经济上是有意义的；以及由博弈和互动模型确定的“全局环境”的不变性，能不能与潜在的经济现实兼容，而博弈规则在现实中将经历内生性变化。要想解决传统数学工具与那些可能表现出永恒新异性现象之间的固有张力，无疑是非常不容易的。

正如我们在前面已经提到过的，第二卷中有几篇文章引入了不那么传统的、基于行为主体的模型。科尔曼、米勒和佩奇讨论了这种建模方法的优点和困难。他们在结论里对这种方法在未来研究中的作用表示了谨慎的乐观。特斯法齐则把自己的文章当作了她所谓的“经济学中‘人工生命’方法的一个例证”，她的文章也阐述了“经济学中仍然有待清理干净的障碍”。说了这么多对这些与“过程与涌现”观点相关的知识论问题，我们可以向读者提供的最好建议也许非常简单：读一读第二卷收录的这些文章，找一找你认为有说服力的东西。

COMPLEXITY

AND THE

ECONOMY

06

再好的经济和社会系统也会被“玩弄”

“压力测试”是防范操控行为的良方

再好的经济系统，也可能存在被他人利用的漏洞。我们可以借鉴工程上的失败模式分析法，对政策设计进行“压力测试”，模型中无须加入任何“剥削”元素。行为主体在做策略选择时，会发现某些行动特别有效，这时“剥削”行为就显露了身影。通过计算机的自动预警，我们就能采取相应的防范措施。

在2008年美国金融危机之后，许多经济学家就已经看得非常清楚了：金融系统，以及其他社会和经济系统，很难避免被各种小团体操纵，以达到他们谋取私利的目的。这样也就自然而然地引出了以下两个问题。第一，给定某种特定的政策框架或拟议中的经济系统，这种操纵是否可以提前预见并预先防止？第二，我们是否可以设计出某种方法（可能是自动的），来测试它们可能会面临失败的模式，以及它们可能

被操纵的程度，从而防范将来可能发生的操纵行为？

剥削行为（**Exploitive Behavior**）　无论是金融系统，还是其他社会和经济系统，都很难避免被各种小团体操纵，以谋取他们的私利。我们把这种行为称为“利用行为”。

本章认为，经济中的剥削行为绝非罕见，而且可以划分为若干特定类别；我们很容易就可以扩展政策研究，以便考察政策被“耍弄”的可能性；经济学需要发展一门强大的、专注于失败模式分析的子学科，就像在结构工程和飞机设计中已经大获成功的失败模式分析的子学科一样。本章完成于2010年，当时我正在IBM公司位于阿尔马登的“健康分析和模拟智慧地球平台”（Smarter Planet Platform for Analysis and Simulation of Health，简称“SPLASH”）工作，并得到了同事们的帮助，在此表示感谢。本章是首次公开发表。

* * *

社会和经济生活中有一个一般规则：给定任何一个系统，总会有人找到一种利用它、剥削它的方法。或者说得更简洁一些，所有系统都会被玩弄。我这样说，并不是因为我愤世嫉俗。恰恰相反，我在这里指出的是一个有普遍意义的观察结果，即任何政府制度、任何法律制度、任何监管制度、任何企业制度、任何选举制度、任何政策组合、任何组织规则、任何国际协议，人们都能够以你意想不到的方式，利用它来谋取自己的私利。2005年，亚利桑那州州长珍妮特·纳波利塔诺（Janet Napolitano）在谈到美国和墨西哥边境的非法移民问题时，模仿非法移民的口气说：“你们把边境墙砌到15米高，我们就会造出16米高的梯子来。”说的就是这样的情况。

能预见到非法移民会造出16米的梯子，也许并没有什么了不起。毕竟，纳波利塔诺只是在表明自己的政治观点。但是，要一般性地预测给定的政策体系中，剥削行为是如何出现却是极具挑战性的。要“剥削”各种系统，方法有很多，其中有些是非常不明显的。但是，我们确实极度

需要预见到各种可能的操纵，因为它们有时会造成灾难性的后果。例如，读者不妨回忆一下俄罗斯1990年从社会主义计划体制向资本主义体制过渡的后果。在那个时期，少数“精英”攫取了国家新释放出来的资产的控制权。或者，想一想加利福尼亚2000年开放能源市场的后果，当时少数能源供应商操纵了市场，损害了整个州的利益。又或者，考虑冰岛2008年的银行体系，一些控制了该国各大银行的所谓金融家，挪用存款人的资产到海外房地产市场上投资，将银行拖入了破产的陷阱。当然，读者很可能仍然对2008年华尔街崩盘记忆犹新，当时对抵押贷款支持证券市场的监管非常宽松，贷款唾手可得，而衍生品则越来越复杂，以至于形成了一个高度不稳定的结构，最终令人震惊地塌陷了。所有这些系统都被操纵了，或者用一个更强、更触目的术语，有些市场被“玩弄”了。现在回顾起来，不难看出，所有这些系统都提供了强大的激励，使得它们非常容易被操纵，而且很多最终都难逃系统崩溃的悲剧。

这就提出了一个不容回避的问题。经济学已经发展得如此精巧复杂，事前又有无数经济学家对拟议中的政策体系进行了细致的研究，怎么还会发生这些经济灾难呢？在前面我提到的那些例子中，确实有若干经济学家提前预见到了未来发生悲剧的可能性并发出了警告。但是这种警告通常并没有什么效果。原因在于，目前大行其道的那种经济学隐含着一种偏见，它会使经济学家选择性失明，因而无法看到经济系统未来被“剥削”的可能性。标准经济理论假定要研究的系统是均衡的，而根据定义，在均衡中没有任何行为主体有动力偏离自己的当前行为。由此而导致的结论只能是，在标准经济学研究的任何一个系统中，侵入性行为或破坏性行为永远不会发生。如果系统可能被侵入，那么某些行为主体将会发起新的行为，并且系统将不能继续处于均衡状态。均衡经济学的基本假设决定了，它的研究目的不是探索各种系统被剥削的原因和结果，因此对于系统是如何失败或被利用的系统性研究，不是这个学科思考的核心问题。[1]

在本章中，我将摒弃均衡假设，并以一个完全不同的非均衡假设为基础来进行分析：任何时候、任何政策体系都会为相关各方提供某些激励，而且这些激励反过来可能会促使各相关方发现一些政策设计者没有想到的、能够用来增进自己私利的途径。在这个假设的基础上，我们就可以探索：“剥削”政策系统的行为是怎么出现的，我们如何对这种行为进行建模和分析，并提前预见到它们或者预先提出警告。

为了便于分析，我把怎样才能预见到各种系统可能被“剥削”这个问

题分成4个子问题依次考察。第一，这种剥削行为的原因是什么，它通常是怎么出现的？第二，考虑给定的特定经济系统或政策，我们如何预测它（们）可能会在什么地方归于失败，以及我们可以从结构工程学等学科中学到一些什么？这些学科都试图预测潜在的失败模式，因此应该可以在这方面帮助我们。第三，我们如何为被玩弄或被利用的系统构建模型，以及我们怎样预测这些模型中的行为主体“发现”“剥削”这些系统的方法？第四，构造人工智能模型来自动预测经济和社会如何被剥削，这种方法在未来的前景如何？当然，我在这里不可能给出这些问题的完全确定的答案，但是我希望本文至少能提供一个辩论的主题。

在展开具体分析之前，我先给出我在下文中将会使用的一些术语的定义。“剥削”（exploitation）这个术语有两层含义。第一层含义是“为了获得利益而使用某种东西”，第二层含义是“自私地或不公平地利用某人或某种情势，而且通常是为了个人私利”。[2]第一层含义适合我们的讨论，但注意这种意义上的“剥削”一词不一定是贬义的。第二层含义也涵盖了我将要谈论的许多情况。“玩弄”（gaming）这个术语，本身就带有一个更加“有害”的含义：它表示人们为了实现自己的目的而操纵性地使用一个系统，通常会背弃他人对他们的信任，并会伤害他人。[3]我也会讨论“政策体系”（policy system）这个术语，意指在给定一整套政策的情况下，随着时间的推移而渐次展开的经济系统、社会系统、军事系统、商业系统或政府系统。例如，奥巴马总统2010年提出的“负担得起的医疗保健制度”，就是一个政策体系。

“剥削”的4种类型

在我们分析如何对剥削行为建模之前，先讨论一下剥削行为的原因和机制，这些“预备知识”无疑是非常有用的。

我们的第一个观察结论是，剥削行为并不罕见。之所以如此，并不是因为有的人天生是自私的，而是因为所有的政策体系，即所有的社会政策，都会产生一些激励。作为对这种激励的反应，由行为主体组成的某些群体会为了自己的利益而采取行动，这种反应往往是意想不到的、违背政策设计者的意图的。这样的例子比比皆是。例如，2003年美军对伊拉克展开军事行动。这是一个军事政策体系。它计划得很好，执行得也很好，但还是引发了叛乱。美国决策者关于伊拉克民众对美国在伊拉克驻军的反应预期不足，这种反应阻碍了美国在伊拉克的目标的实现。

又如，美国在林登·约翰逊担任总统期间于1965年启动的医疗保险体系，目的是为老年人提供医疗保健、支付医疗服务费、补偿医院和医生的治疗费用。参加这个计划的医院和医生对它的反应是，购买昂贵的设备和提供不必要的医疗服务，因此在该计划推出后的短短5年之内，它的成本就几乎增加了3倍。12年后，美国向市场开放了医疗保健服务，目的是为了引入竞争和降低成本。但是事与愿违，结果却产生了这样一个系统：每个关键的参与者都能找到某种方法，来利用这个系统谋取自己的利益，从而有害于整个系统。马哈将它描述为一个“霍布斯式市场”，他这样写道：“医疗保健行业的所有博弈参与人都在相互对抗：医院与医院、医生与医院、医生与医生、医院与保险公司、保险公司与医院、保险公司与保险公司、保险公司与药物制造商、药物制造商与药物制造商。”

上面举的这些例子都是一些“大规模的剥削行为”，其实剥削发生在所有尺度上。众所周知，公寓楼的管理人员，会溜进他们竞争对手的公寓楼窥探虚实，并在网上发布竞争对手的负面评价，以提高自己的竞争地位。无论剥削的规模如何，其发生频率之高已经足以让我们考虑暂停实施一些社会政策，而我们并没有仔细设想过这些政策被相关参与人利用的可能性。这个事实还警告我们，在接受那些旨在展示某个政策体系的“美好愿景”的经济模型的结果之前，一定要三思而后行。事实上，我们应该时刻保持警觉：在对它们的内在假设进行恰当质疑之前，任何政策模型的结果都不能贸然接受。

那么，我们应该如何质疑政策体系的结果呢？我在上面给出的例子看起来是分散和独特的，要想从它们那里得出一般的洞见似乎并不容易。如果我们可以确定剥削行为的“通用类别”，即剥削的“标准技巧”，或者如果我们能够总结出各种情况下反复出现的剥削行为模式或激励模式，那么无疑会非常有用。换句话说，如果我们能够在经济学中开创一个用来评估政策体系的“失败模式分析”传统，那么将会非常有用。在其他一些学科中，这样的传统很早以前就存在了。在那些学科中，当生命或安全或面临危险时，就会进行这种“失败模式分析”。例如，在工程学中，失败模式分析或故障模式分析研究的是：结构在过去遭受失败的方式，或未来可能失败或无法按预期发挥作用的方式。又如，在预防医学和疾病控制学中，要研究疾病、死亡和流行病的原因，探索未来预防的方法。这些研究不仅要探析过去的失败，而且要建立一个有序的知识体系，以帮助预防未来可能出现的失败或崩溃。

要想构建一门专门研究政策体系失败模式的经济学子学科，无疑是一个非常艰巨的任务。这种努力当然是非常有价值的，但是超出了本文的范围。我们现在能够做的是，细致地思考一下这门子学科怎样才有可能成功地建立起来。而一个很好的开端就是，考察各种系统在过去是如何被剥削或被玩弄的，并对这些发生的事件进行分类。下面，我将讨论4种类型的剥削，并用对应的原因命名。

1．利用不对称信息。

在许多社会系统中，系统各相关方可以访问不同的信息。经常发生的一种情况是，某一方根据自己对可得信息的理解提供某种服务或某种机会，然后另一方以自己所拥有的更详尽的信息为基础做出回应、采取行动，并依靠其优势信息利用系统从中获益。在金融行业和营销行业中特别容易出现这种行为。因为在这些行业中，推广自己产品的那一方非常了解他们正在推广的产品，而另一方潜在的投资者或客户则不然。2007年，高盛创设了一个抵押贷款债券资产包，向它的客户兜售。但它又请了一个著名的对冲基金经理约翰·保尔森（John Paulson），来为这个资产包选择债券。私底下，保尔森认为这个产品将会变得一文不值，于是以它为标的进行对赌。通过购买可以避免这种金融工具带来损失的保险，保尔森大发横财，高盛也小赚了一笔，但是投资者却损失了10亿多美元。这个资产包，即与次级住宅抵押贷款支持证券相关联的综合性抵押债务，是一个非常复杂的“套餐”，其设计者高盛和保尔森对它的前景非常了解，但是他们的客户却无法搞清楚。

医疗保险行业也很容易出现信息不对称问题。医生和患者都比保险公司或为他们支付医疗费用的政府机构更了解疾病情况和何种治疗才是适当的。2006年，马萨诸塞州立法推进个人医疗保险计划。这个计划一开始似乎是有效的，但是短短几个月后，保险公司就发现他们亏钱了。原因在于，正如苏德曼（Suderman）所指出的那样，“成千上万个消费者都在‘猎杀’马萨诸塞州2006年健康保险计划，当他们需要支付昂贵的医疗护理费用，如生殖治疗和膝盖手术的费用时就参保，然后就迅速弃保”。这种行为不是非法的，也不是不道德的，但它确实是掠夺性的，是一种剥削。

2．“裁剪”行为以符合特定标准。

第二种类型的剥削，在这里也许称之为“操纵”更好一些，发生在必

须根据严格的评估标准对行为主体的行为进行评判、监控或测度时。行为主体会“裁剪自己的行为”，以便符合这些标准的狭隘的条文，而不去考虑制定这些标准的更广泛的意图。换句话说，行为主体“玩弄”了标准。在2008年金融危机之前，金融评级机构如穆迪或标准普尔等，多年来一直在对投资银行创设的各种金融工具所固有的风险进行评估。但是在金融危机爆发前几年，在一个以增加透明度和信任感为名的行动中，它们将自己的评级模型提供给了华尔街各投资银行。对此，摩根森（Morgenson）指出：“华尔街的投资银行很快就学会了如何篡改这些模型，如改变一两个小小的输入项，然后得到更好的评级结果。也就是说，他们学会了如何利用评级机构的模型作弊，目的是保证在把质量更低的债券掺入投资组合后，仍然能够得到高评级，于是他们就可以把这些本来很可能卖不出去的垃圾卖出去了。”[4]

玩弄绩效标准的人绝对不限于华尔街那帮银行家。在一切需要评判性能或绩效的系统中，都存在这种行为：是否符合法律、能否通过教育考试[5]、是否遵循人权标准、有无遵守环境标准、是否达到接受资助标准、工厂产量是否达产、金融会计是否公允、税务报告是否合规、行政官僚业绩是否出众、政府治理的绩效能不能令人满意等，涉及的标准都可能被“玩弄”。在所有这些情况下，受到监管或被关注的当事方会调整自己的行为，做出一副在规定的绩效措施下“表现良好”的样子，但是他们的实际行为却千差万别：既可能令人满意，也可能应受谴责。事实上，以政府治理绩效为例，人们早就总结出了这种“剥削”的两个“定律”。第一个是“坎贝尔定律”（Campbell's law）：“社会决策越是频繁地使用任何量化的社会指标，招致腐败的可能性就越大，也就越容易扭曲和腐化它原本打算监管的社会过程。”第二个是“古德哈特定律”（Goodhart's law）：“任何观察到的统计规律性，只要将它用于控制目的，就必定会失效。”[6]这两个定律都适用于政府行为。不过，我本人更喜欢一个更宽泛的表述：任何绩效标准，都会被最大限度地利用，从而丧失价值。

3．获得系统的部分控制权。

第三种类型的剥削，发生在当由一些行为主体组成的小团伙控制了系统的部分重要资源，并将其用于满足自身的目的时。经济系统中的这种剥削，类似于计算机系统中病毒对计算资源的剥削。这个小团伙实际上接管了一部分系统，并且利用它来为自己谋取利益。

在金融部门，我们可以观察到这种类型的剥削的许多例子。在2008年金融体系崩溃前的几年里，保险巨头美国保险集团（AIG）内部，一个小团伙，即该公司金融产品部，设法在实际上控制和管理了公司的大部分资产和风险负担，并开始大量投资于信用违约互换。这个小团伙的成员以高额薪酬的形式获得了巨大的利益，即他们可以获得他们所“赚得”利润的1/3，但是投资最终失败了，这令美国保险集团遭受了灭顶之灾。冰岛也发生了类似的一连串事件，一小部分企业家利用贷款买下该国银行资产的控制权，并将银行资产投资于国际房地产市场和相关的衍生品。这些投资失败了，冰岛的银行体系垮塌了，它们客户的存款都成了殉葬品。

4．以政策设计者意图之外的方式利用系统的元素。

当行为主体利用系统自身去操纵系统时，还会出现另一种类型的剥削。一个例子是，利用网站的评级系统去操纵对他人的评级。又如，有关利益方会设法去搜寻某个规则，然后使它成为系统的一个漏洞，用来证明他们所做的超出了系统设计者意图之外的行为的合理性。这样一来，这条规则就会蜕变成一个管道，让金钱或能量滚滚流过，并对整个系统造成不利影响。在20世纪70年代初，阿拉伯对石油实行禁运之后，美国国会为机动车辆制定了燃油经济性标准，不过对商用轻型卡车的要求比对乘用车的要求更宽松一些，这也是可以理解的。但是，底特律的汽车业人士在进行了大量的游说活动后，操纵了国会在一个适当的时机宣布运动型多用途车（SUV）是轻型卡车。这样，大量汽车就轻松地“驶过”了这个“轻型卡车漏洞”，导致美国的高速公路上塞满了SUV。结果是，在1988年和2005年之间，美国的平均燃料经济性实际上反而下降了。这绝对不是这项能源政策设计者的本意。

上面描述的只是4个类型，它们绝对不是全部。毫无疑问，各种系统被“玩弄”的方式还有很多。但是，仅仅是这些，就已经足以让我们对剥削类型有一个深刻的印象了。它们告诉我们，系统中的剥削行为绝非罕见，它们简直成了一种狂欢。

“压力测试”的作用

对于某些政策体系来说，它们可能遭受的剥削，显然属于上面给出的4大类型之一。但是，许多其他政策体系遭受剥削的模式却不是这么

显而易见的。给定一个政策体系，人们就会形成一个心理模型或分析框架，用来预期它将如何运行。我们现在希望可以预计政策体系在现实生活中的哪些方面可能会被利用。那么，我们应该怎样着手去做呢？在一个特定的经济环境下，我们将如何进行失败模式分析？对于这些问题，没有固定的答案，但是我们应该可以借鉴工程学进行失败模式分析的方法和步骤。

显而易见，第一步是掌握关于以往类似系统的失败的知识。有了我在上面描述过的那4种剥削类型，我们可以说至少已经开了一个头。飞机设计师可以从过去的经验中了解到，故障（他们称之为“异常状况”）发生的一般原因包括：疲劳性失效、爆炸性减压、火灾、爆炸、发动机损毁，等等。依此类推，正如我在前面所说的，我们必须对过去政策体系如何被利用的失败模式进行分析。

第二步，我们可以观察到，结构的分崩离析始于比其“总体设计”更微观的水平上的失败。在工程设计中，崩溃往往不是因为总体结构整个散架了，而是因为在应力作用下，部分组件中出现了细的裂纹，或者因为某些组件失效了，这些故障被传播到了更高的层次上，从而最终导致整个系统走向失败。这就意味着，对于我们正在研究的任何系统来说，剥削行为通常不会发生在整个系统的层面上。毕竟，剥削行为是个体，或者是由若干个体组成的小团伙创造的或“发明的”。如果我们想要搞清楚对系统的操纵是怎样发生的，我们就必须细致地探究行为主体所拥有的各种可选项和行为可能性。

第三步，通过类推可知，我们可以在制度中找出应力特别高的那些地方，集中注意力观察它们。在社会系统中，这些地方往往会给予行为主体强烈激励，使他们做出与被规定的行为不同的行为。在分析模型中，这类行为的关节点通常被表示为一种“费率”，如个人购买健康保险的比率；或者被表示为一个简单规则，如收入超过了X美元，并且年龄超过了Y岁，就得购买健康保险。建模者必须追问简单的费率或规则，是不是真的能够保证给定行为主体面对的激励。当然，很多时候它们是不能保证的。

上面所有这一切都表明，如果我们对社会系统有一个构想，并且有一个用来分析社会系统的解析模型，那么我们就可以对它进行“压力测试”，以便识别那些可能会强烈地诱使行为主体做出不同于设计者预想的行为的激励。这些应力很大的地方，可能就是行为主体有权利去影响其他相关方福利的地方，如他们可以向富裕的房地产开发商颁发建筑许可

证，但是我们通常假设他们会做出公正的决定。有些时候，行为主体可以通过牺牲一定程度的性能或安全性来谋取利益，如他们决定降低飞机维护的频率，但是我们通常会假定一切都符合给定的标准。另外，行为主体有时会利用内幕信息谋利，如他们知道公司的未来计划，但是我们通常会假定他们不会进行内幕交易。

接下来，我们就可以根据我们所了解的关于行为主体的具体激励和他们掌握的信息知识，来建构行为主体的“行动可能性”了。换句话说，我们可以为行为主体构造一个详细的策略选择侦查。这里的关键词是“详细的”：所包含的选项或可能性，是由观察该系统的分析者的想象和经验决定的，它们源于现实世界。我们需要把它们认真地、详细地描述出来。我们之所以必须尽可能地去了解一切与行为主体据以决定自己行为的微妙信息和可能机会，原因也就在这里。

一旦我们搞清楚了剥削可能发生在什么地方、又是如何发生的，我们就可以打破关于这个政策体系的整体经济模型，并插入一个模块用来“注入”我们想到的那种类型的行为。这样一来，我们的脑海中就有了一种特定类型的剥削者的形象，以及关于这种剥削的工作模型。于是，我们就可以用它来研究不同的策略性行为主体对整个系统的行为所造成影响的差异了。有的时候，它们只会造成细微的差别，因为策略性行为也许不会在其自身的领域之外造成太大的影响。但是另外一些时候，它们却可能造成重大的影响。在某些情况下，它们甚至会导致所嵌入的结构塌陷。在这里，我们正在寻找的是政策体系中的薄弱点，以及特定行为可能对系统造成的后果。我们还一定要记住，这种测试不能操之过急。在工程学中，当出现了一个重要的新设计时，通常需要几个月，甚至几年的时间来精心检验、调试和重新设计，尤其是当它事关公共安全大局时。我们没有任何理由认为，人们在制定经济政策和社会政策时，反而不用强调它们的后果是否安全。

我在上面概述的这种方法，适用于只有一个设计者或一个设计小组的情形，他或他们独自去发现给定的政策或经济系统的缺陷。如果有多个设计者或小组齐头并进地工作，都致力于探测某个模型的弱点，那么整个过程无疑可以加快很多。对于拟议中的政策体系，不妨想象一套新的医疗保健政策，或一组新的金融管制条例，一旦有了一个“工作模型”，我们就可以通过它诱导出一个剥削该体系的“策略”模型。更具体地说，我们将先给出一个整体模拟模型或总体政策环境，然后任由外部参与者提出自己的策略来剥削它。类似的做法由来已久。早在几十年前，

罗伯特·阿克塞尔罗德（Robert　Axelrod）就举办过一系列后来变得非常著名的囚徒困境博弈锦标赛：让不同的策略在重复的囚徒困境博弈中相互竞争。在一个更具一般性的系统中，参与者需要对该系统进行彻底的研究，将它的数之不尽的激励一一识别出来，占据那些存在剥削机会的地方，然后形成相应的模型。

类似的情况也经常发生在加密系统的beta测试中。例如，假设美国海军开发了一种新的加密方案，它会邀请一群人来参加测试，即看看他们能否破解该加密方案。如果他们不能，那么这个计划就可以继续推进。在这种情况下，必须保证测试人员来自外部，这是非常重要的。又如，就像施奈尔（Schneier）所指出的："就拿互联网IP安全协议来说吧。它是由一个委员会公开设计的，从一开始就受到相当多的公众监督……海军研究实验室的密码学者最近发现了一个小小的的缺陷。这项工作继续公开进行着，任何一个有兴趣的人都可以参与测试。另一方面，微软公司也开发了自己的点对点隧道协议（PPTP），其功能基本相同：他们发明了自己的身份验证协议、他们自己的哈希函数，以及他们自己的密钥生成算法。这些发明每一个都有严重缺陷……但是，因为他们是在内部完成所有这些工作的，没有人知道他们的PPTP是薄弱的。"

用"涌现"发现新行为

在上一节中，我们讨论了探测政策体系失败的可能性的一般方法。现在，我想收缩范围，集中讨论用于探测政策体系失败的基于计算机的模型，通常来说是计算机仿真模型。但是，我们立即就会面临一个困难，那就是大多数基于计算机的模型都很难容纳新异行为：它们往往采用微分方程组、马尔可夫状态或其他建模架构，并假设了事先预置或嵌入在系统内部的固定类别的行为，因此要修改它们，让它们"召唤出"不可预见的行为，即可能出现的"高16米的梯子"并不容易。

但是，我们还是可以向前推进的。很显然，正如我在前面已经提到过的，一个方法是我们向计算机仿真模型中的系统注入"事前推想"的剥削行为。虽然这样做可以让模型保持开放并加入更多的细节，但是我们希望有更一般的方法，就是让我们的仿真模型主动"发现"不可预见的新行为。这无疑更具挑战性。需要注意的是，我们现在实际上要探究的是，系统中的新行为如何在行为主体的发现或学习过程中涌现出来，而且我们运用基于行为主体的建模方法处理涌现问题，已经有相当多的经

验了。因此，我们有理由相信，我们确实可以修改某个模拟模型，来允许行为主体“发现”操纵行为。

接下来，我先用一个真实世界的例子来说明这一点，即我在前面提到过的马萨诸塞州的医疗保险计划。我们现在没有关于这个政策体系的现成的模拟模型，所以我们不得不自己构建一个。因为我们感兴趣的不在于描述所有行为的社会性方面的细节，而在于如何模拟剥削行为，所以我们可以构建一个简单和典型的模型。我们将逐步推进，依次开发出这个模型的一系列版本，来刻画我们感兴趣的剥削行为。

第一步，我们构建一个基本的医疗保险模型。我在进行计算机仿真时使用的是NetLogo，这是一个非常方便的构建基于行为主体的模型平台。在这个模型中，有N个行为主体，但在NetLogo平台上，通常为100～1 000人，他们单独地、随机地产生各种各样的医疗费用，这些费用可能源于疾病、医院护理、外科手术、事故等。初始状态是，他们自己承担所有这些费用。在这个模型中，个体健康成本的分布是一致的、固定的，并且对于所有行为主体都是相同的，如我们可以假设所有人的年龄都相同。人们的收入是固定且相同的，他们的医疗消费c等于他们的保健成本。对于医疗消费带来的效用，我假设了一个凹的效用函数$U(c)=c^{1/2}$。再假设所有人都是风险厌恶的，系统中只有一家保险公司。一开始，保险公司不提供任何保单（或医疗保险政策），而只是在固定时期收集精算数据：它能够获得所有人的保健成本数据，并会使用这些数据来计算每人每期的平均保健成本。一旦它得出了一个足够准确的估计，它就会发售一份自愿的健康保险单。该保单的成本是这样设定的：在“公平”的保健成本，即保险公司估计的每人每期预期成本的基础上，加上m%的加成，用以支付管理费用。通过计算机仿真实验，我们观察到在这样一个模型中，当上述成本加成足够低（$m<23.3\%$）时，人们就会发现，购买保险的效用更高，而且他们确实也会购买保险。不然的话，他们就不会购买。现在，我们已经有了一个简单的基于行为主体的保险模型。

第二步，我们要对马萨诸塞州的保险计划法令及其后果进行建模。这里出现的问题的核心在于，有这么一类人，他们相信或发现，不参加这家公司的保险计划对自己更有利，这就是说，他们宁愿因自己的不参保行为而向政府缴纳罚款。要想让我们上面那个模型容纳这种类型的行为主体，有好几种方法。例如，我们可以假设，系统中有一类人面临承担健康成本的风险很小。但是最简单的方法是，假设一部分人，不妨假

设50%的人不是风险厌恶型的，即他们拥有一个线性效用函数$U(c)=c$。这样的话，他们就会发现，只要向政府缴纳的罚款低于那家保险公司的保险加价（就马萨诸塞州的案例而言，确实如此），那么就向政府缴纳罚款。我们运行了这个模型，结果发现一半的人参加了保险，另一半人则没有。这个结果并不令人奇怪。

第三步，让我们构建关于剥削行为的模型。现在，我们要允许所有人都可以预先知晓某些类型的医疗保健如肩部手术或物理治疗的费用。为此，我们在模型中内置了，或者说，“注入”了这些行为，并让它们在每一期刚开始时就可以被预见到。同时，我们还赋予这一类人选择权，允许他们在某一期投保，并可以在下一期取消保险。对这个模型的计算机仿真结果表明，已经投了保的那50%人不受影响，他们无论怎样都会继续支付保险金。但是，本来没有投保的那些人的行为则会受到影响。我们发现，他们会根据这种保险是否适合自己而随时决定加入或退出保险。从某种意义上说，他们正在为他们预先知道的某种结果购买保险，而保险公司则不知道这种结果。因此，他们是“玩弄”了这个系统。图6-1显示了保险公司在这种情况下的利润。很显然，它的利润出现了暴跌。

第四步，作为我们的建模过程的最后一个阶段，我们可以假设系统会对这种行为做出反应。这是符合现实的。一方面，政府可以提高对不参加保险计划的罚款。我们发现，如果政府真的这样做了，那么每个人都会参加保险，系统也就恢复了常态。另一方面，保险公司也可以做出反应，即增加强制性的投保期限。保险公司的强制性投保期限一旦变得足够长，我们就会看到系统再次恢复正常。

我之所以要分步详细介绍构建模型的每个阶段，是因为这确实是一种非常便利的方法。先建立基本模型，然后打破，将做出策略性反应的那类行为主体区分出来，并允许他们这样去做，同时考虑其他行为主体的“自然”反应。当模型最终完成时，当然会遍历所有这些动态步骤。

到目前为止，我们已经阐明，对于特定的政策体系，我们可以采用一个计算机模型，也可以像上面这个例子一样构造一个计算机仿真模型，并通过引入预想得到的“剥削”行为，以及对这种行为的反应对模型进行完善。但是，在现实世界中，剥削行为是自发地涌现出来的。

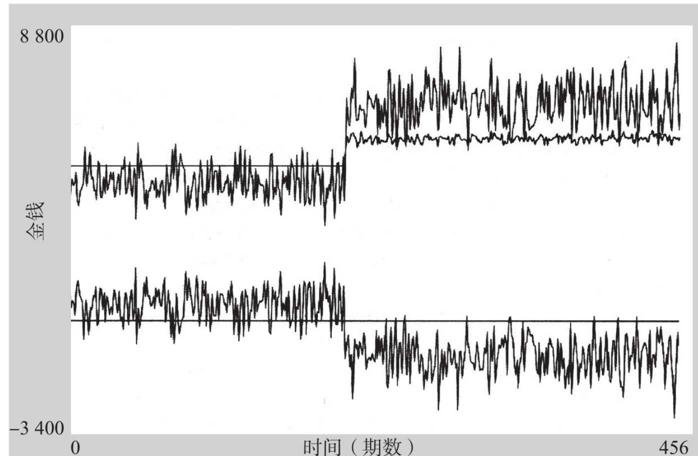


图**6-1**　保险公司的收入费用支出和利润变化趋势

图的上半部显示了保险公司的收入受到的影响。保险公司的收入增加了，因为它的投保人增加了。保险公司的收入是较平滑的那条线，上面的锯齿线是费用。图的下半部则显示了保险公司的利润，即下面的锯齿线。不难看出，利润现在已经下降到了0以下。

这就是说，这种行为是在政策体系存在的过程中出现的。如果我们更仔细地观察现实生活中所发生的事情，就会发现相关参与者会注意到他们可以利用的某些选项，他们会学会，或者是突然发现，可以通过某些行为来获利。因此接下来，我们先看看怎样才能将“注意”和“发现”内置于我们的模型中。为了保证行文简洁，在这里我只概略地说明如何做到这一点。[7]

首先，“注意”当然是相当直接的。要允许我们的行为主体“注意到”某些东西非常简单，如最近发生了哪些事件，哪些选项是可行的，只要让他们在信息集的这些部分出现时，能够“知晓”它们就可以了。

不过，我们还需要对“发现”建模。为了这个目的，我们允许行为主体根据他们的信息生成和尝试各种潜在的行动或策略。有很多方法可以实现这一点。行为主体可以随机地生成各种类型的状态依存的动作或规则：如果系统满足一定条件K，那么就执行策略G。或者，他们可以通过对以前采取过的行之有效的行动进行重新组合，来构造新行动：如果条件K和P为真，那么就执行策略F。又或者，他们也可以生成一个可能行动的集合：如果这一期预先知道医疗保健费用超过了k美元，那么就购买保险。这里的k可以锚定在多种不同的水平上。进一步，我们假设行为主体能够记住这些潜在的策略（可能有许多），并能够监控和记录每种策略的推定结果，这样一来，随着时间的推移，他们就能学会哪些策略在哪些情况下是有效的。然后，他们就可以随时使用或执行他们认为

最有效的策略，并丢弃被证明无效的策略。

这样设计模型，可以为我们带来我们想要的效果。如果被“监控”到的某种随机生成的策略被证明是特别有效的，那么行为主体就会特别快速地“发现”这种策略。从局外观察者的角度来看，这看上去就好像是这种策略在突然之间就“绽放”开了，这也就是说，它突然涌现出来并产生了效果。但是，行为主体是以归纳的方法去探测所处的系统，看到底什么东西是有效的，从而随机“发现”了利用系统的有效策略。由此，“剥削”也就涌现出来了。

在上面的例子中，我已经描述了一个相当简单的模型，并概略地阐明了对可能剥削行为的涌现进行建模的方法。很显然，我们可以从很多角度来进一步深入阐述。[8]

但是在这里，我想再强调一下最重要的一点。当我们想要对操纵系统的行为进行建模时，并不需要往基于行为主体的仿真模型里加入任何用来代表“阴谋诡计”或“剥削思想”的因素。行为主体只需要根据关于系统及可用选项的特定信息，决定可以采取的行动。有时，他们会发现有些行动特别有效，在这种情况下，我们所称的“剥削”就出现了。要对此建模，只需运用“标准方法”，这种方法在基于行为主体的建模中早就存在了。

自动预警

但是，现在还没有到最后下结论的时候。在上一节中，我们看到，如果我们想要通过计算来“发现”剥削行为，还需要指定某个特定类型的行为主体，他们会进行探索。在未来，我们希望通过计算能够自动“发现”我们没有想到过的、更广泛多样的可能剥削行为，并且对它们进行测试，从而提前预测到可能的操纵。如果真的能够做到这一点，那么无疑理想得多。

前景究竟如何？针对美国–墨西哥边境的偷渡，基于计算机的模型将会给出什么预测（如它是否能够想象到16米高的梯子）？基于计算机的模型在得到了我们的提示后，很容易就能做到这些。我们为计算机仿真程度提供不同的“梯子高度”，如8米、12米、20米等，这样计算机仿真程序就会“学会”：16米高的梯子是有效的。但是，这其实是在作弊。

就“偷渡攻防战”这种情形而言，我们真正希望做到的是，计算机仿

真程序能够在完全没有人为提示的情况下自主地“思考”：当边境上竖起了防止偷渡的高墙时，怎么办？我们希望计算机仿真程序能够在没有任何内置提示的条件下，“发现”梯子这一类工具，或者发明其他合理的方式来战胜障碍。而要想做到这一点，计算机仿真程序就需要有关世界的知识，这种知识必定是一种深刻的知识。为此，计算机仿真程序必须成为一个一般意义上的智能体，知道世界的各种可能性、知道什么是可用的、知道一般来说“外面会有什么”。换句话说，它需要某种类似于我们人类智力的东西。不难感觉到，这里有一些人工智能的味道。是的，我们其实是在要求得到一台“发明机”，即一台能够意识到自己身处的世界、能够在概念上将一系列构件组合到一起去解决一般性问题的机器。从这个角度来看，这个问题属于“人类认为不难，但对机器来说却很困难的计算问题”，即所谓的“人工智能完备性”问题，诸如阅读和理解文本、解释语言、翻译语言、识别视觉对象、下棋、裁判司法案件等问题，均属此类。考虑到我们关注的焦点，我们还可以加上一项：对解的想象。

这里的问题绝对不仅是一个概念性的问题，还是一个实践问题。如果能够认识到这一点，那就是一个令人高兴的良好开端。我们能够“教会”计算机识别不同情境，并让它建立一个庞大的、通用性的实用知识库。众所周知，早在2010年，IBM就“教会”了一台计算机，并让它成功地参加了电视智力竞赛节目《危险边缘》，它就是通过建立一个庞大的知识库来实现这一点的。因此有朝一日，计算机也会拥有一个庞大的、关于以往各种情形的，以及这些情形如何被“剥削”的语义性知识的“图书馆”。有了这样一个“图书馆”，计算机就可以“识别”出类似的情形，并利用它们来为当前的目的服务。这并不是一个遥远的预言。例如，在2003年，如果在美军向伊拉克发动进攻之前，已经进行过了计算机仿真实验，那么就应该能模拟出或计算过历史上的入侵，以及随着这种入侵而来的叛乱，还应该能就伊拉克的可能未来提出警告。它会预料到种种“涌现”出来的行为。在未来，计算机模拟应该可以很好地融入历史，找到类似的东西，如找到作为对高墙反应的梯子，并将这些东西呈现给我们。尽管从概念上看，这是完全可行的，但是我认为要充分利用这种类型的实用机器智能，可能还需要几十年的努力。

补上“失败模式反思”这一课

在过去100年或更长的时间里，经济学理论在稳定宏观经济预期、制定国际贸易政策、监管货币体系、推行中央银行制度和执行反托拉斯

政策等方面发挥了重要作用。但是，它未能防止金融崩溃和经济危机，而绝大部分金融崩溃和经济危机都是由剥削行为造成的。在我们的时代，这几乎可以说是一个异数。航空安全、建筑安全、地质安全、食品和药物安全、疾病安全、手术安全等，所有这些在过去的50年中一直在稳步地得到改进。相比之下，“经济安全”在过去50年里却没有什么改变。如果说有什么改变的话，那也只能说它变得更糟了。

许多经济学家（包括我自己）也许会说，对自由市场自我调节能力的盲目信心是造成这种情况的一个很重要的原因。但是，经济学理论本身的一个弱点，即缺乏一种能够在政策实施之前找到可能的失败模式的系统方法，也难辞其咎。失败模式研究从来不是经济学这门学科的核心，因为经济学坚持均衡分析，假定系统会快速地向一个没有任何行为主体有动机偏离其当前行为的地方收敛，并且会稳定在那里，因此剥削行为不可能发生。基于这种认识，我们经济学家的普遍倾向是，设计政策并对其结果进行一些模拟，但是不会充分地探讨行为假设的稳健性，不能将那些因系统性的剥削而可能失败的地方识别出来。

因此我认为，现在是对我们的思想做出修正的时候了。设计一个政策体系并简单地分析是远远不够的，即便是相当细致深入地模拟政策结果也是不够的。我们不能把社会系统和经济系统视为一组没有改变动机的行为，而必须把它们视为一种总会引发进一步的行为、诱致进一步的策略、导致系统性改变的激励网络。我们需要仿效结构工程学、流行病学或加密科学等学科中的做法，预测我们所研究的系统中可能被剥削的地方。我们需要对我们的政策设计进行压力测试，来找出它们的弱点，看看我们能否“打破”它们。数十年来，工程学一直坚持失败模式分析，从而使我们拥有了能飞行数百万公里的飞机、不会在地震中倒塌的高层建筑物……如果将这种失败模式分析应用于政策制定的世界，那么我们将更有可能得到我们所可能的经济和社会结果，从而避免世界上的诸多痛苦。

COMPLEXITY

AND THE

ECONOMY

07

技术究竟是如何进化的

在实验室中观察到的组合进化

技术的进化机制，就是创造新的组合并选择那些有效的组合。所有技术都是从已经存在的技术中被创造出来的。复杂技术是涌现出来的！不过，这是通过先创造简单的、作为构件的技术来实现的。正如我们所预料的那样，只有少数技术被证明是创造“子孙后代”技术的关键构件。

新技术是由已经存在的技术构建而成，或者说组合而成的，而且新技术自身反过来又会作为构件，用于创造更进一步的新技术。通过这种方式，技术，即社会可用的设备和方法的集合，从自身构建出了自身。我将技术进化的这种机制，即创造新的组合并选择那些有效的组合，称为“组合进化”。在2009年出版的《技术的本质》一书中，我对这种进化机制进行了详细的描述。这种进化机制不同于达尔文的进化机制，后者

依赖于因变化和选择而发生的增量变化的不断积累。

组合进化（**Combinatorial Evolution**） 是技术的一种进化机制。所有技术都是从已经存在的技术中被创造出来的，已有技术的组合使新技术成为可能。

那么，能不能在“实验室”中证明技术的“组合进化”呢？2005年，沃尔夫冈·波拉克和我设计了一个用来检验“组合进化”的计算机算法。从一组原始逻辑电路，即与非门开始，让它们随机地组合成其他更复杂的逻辑电路。如果所得到的某个电路能够完成逻辑上有用的事情，就保留下来并封装好，然后作为一个构件加入到原来的“电路之汤”中，用于进一步的随机组合。在实验中，我们发现，随着时间的推移，这种连续集成的过程越来越复杂，我们最终得到了大量复杂的加法器、比较器和异或器。组合进化确实可以创造复杂的技术，但这是通过先创建更简单的、作为构件的技术来实现的。我们的结果响应了生物学的发现：复杂的特征只有在较简单的特征首先被进化女神青睐并成了“垫脚石”之后，才有可能被创造出来。我们还发现，有证据表明，由此而产生的技术集合存在于自组织临界状态下。总之，我们的结果证明了这种进化形式的强大力量。本章是我和沃尔夫冈·波拉克合撰的一篇文章，最初发表于2006年的《复杂性》杂志上。

*　*　*

技术永远不可能无中生有地凭空出现。新技术是用以前就存在的组分构建的，或者说组合成的。反过来，这些新技术本身又会成为可能的组分，用于构建更进一步的新技术。[1]在这个意义上，技术，即某个社会或文化可用的机械设备和方法手段的集合，从自身构建出了自身。[2]1912年，人们将原先就存在的三极真空管与其他一些现存的电路组件组合起来，发明了放大器。放大器使得可以产生纯正弦波的振荡器的出现成为可能。这些新组件与其他组件的组合，又使得可以改变信号频率的外差混频器的出现成为可能。这些新组件与其他标准组件的进一步组合，又使得连续波无线电发射机和接收机成为可能。无线电发射机和接

收机与其他元件的组合，又使得无线电广播成为可能。

在“集体”意义上，技术形成了一个由多个元素构成的集合或网络，在这个集合中，现有元素不断地构建出新元素。[3]随着时间的推移，这个集合不断地通过组合简单元素来构造更复杂的元素，而且可以用很少的构件元素创造许多元素。这种进化不仅是由先前技术的可用性驱动的，它还受到了人类的需求以及技术本身所带来的需求的驱动。特定的需求，如在人类历史上，对食物、交通、疾病治疗、田地灌溉和排水等方面的需求，总是先通过简单的技术来满足，然后用更复杂的技术替代这些简单的技术。被替换的技术（想想马车运输）变得过时后，还会使得依赖于它们的其他技术（想想马车车厢制造业和打铁业）也变得过时了。因此，新元素不仅会使网络扩增，同时还会产生熊彼特所称的“创造性破坏”。当然，所有这一切都是通过经济能动性，以及工程师、科学家和开发人员的人类能动性实现的。在这里，我们可以将经济视为一个安排技术去满足需求的组织结构。

通过考察人类历史上各项技术一步步地被构建出来的过程，可以探索技术的进化规律。但是，在本章中，我们采取了不同的研究策略。我们在计算机内构建了一个简单的人工世界，来模拟技术被构建的过程。在这个人工世界中，技术，即被构建的元素是逻辑电路。逻辑电路的优点是，我们可以精确地描述它们的功能，而且可以将它们通过相互组合而形成的简单规则界定清楚。我们假设，在我们构建的这个人工世界中，存在一些“逻辑”需求，如对执行异或功能或进行三位数加法能力的需求，而且这些需求可以通过适当的逻辑电路得到满足。当然，前提是它们能够被创造出来。在我们的大多数实验中，原始电路就是简单的与非电路。因此，从原始技术开始通过对现有的电路进行随机组合并测试由此而得到的新电路是否满足某些需求，就可以构建出新电路，即新技术。如果一个电路被证明是有用的，即能够比它的竞争对手更好地满足某种需求，那么它就会取代以前被用来满足某种需求的电路。然后，它会被添加到当前的活跃技术集合中去，并且成为可以用于构建另外电路的元件。由此，有用的元素不断地被添加到当前活跃的技术集合中，而那些被其他元素“击败”的元素则被舍弃。也就是说，技术集合以一种自展向上的方式，先创造能够满足简单需求的简单技术，然后再利用简单技术创造能够满足更复杂需求的更复杂技术。

在这里，我们要提出如下问题。在我们的人工系统中，技术进化的特性有哪些？技术网络是通过怎样一些步骤扩展进化的？某些技术是否

会成为使能技术（enabling technology），即在未来进一步的组合中发挥多种多样作用的技术，如矿石冶炼术或晶体管技术，从而使得生成更多技术的有用性的分布是高度偏斜的？我们能不能在这个人工世界中，观察到熊彼特所说的破坏性的创新风暴？如果我们从一项原始技术开始，我们这个人工系统可以创造出满足复杂需求的元素组合。或者换句话说，我们这个人工系统是不是可以从一个原始的电路进化为我们所需要的电路，比如4位到8位的加法器？需要注意的是，我们的兴趣在于探究复杂的技术进化过程。我们不关注如何制造能够完成布尔函数运算的有效的逻辑电路，那是一个工程学问题，而且已经得到解决了。

先来看上面最后一个问题。在现实生活中，复杂技术的创造既源于更简单技术的存在，也源于使这些更简单的构件得以出现的特定的实际需求。例如，如果没有电子放大器和电子波发生器等构件，如果不存在使这些更简单的构件得以出现的需求，那么雷达的发明就是不可能的。因此，我们不应该预期复杂的电路会在这样的情况下被创造出来，即不存在中间元素，也不存在导致这些中间元素出现的更简单的中间需求。

类似地，在生物学中我们也观察到，如果没有“中间结构”及利用这些中间结构的“需求”，复杂的生物器官，如人眼，是不可能出现的。

我们发现，在我们的人工系统中，技术集合可以从极端简单的电路自展地生成复杂得令人惊奇的电路。我们还发现，正如我们所预期的，被创造出来的大多数技术创建都没有成为特别有用的构件，但是有些技术则被证明是创建“子孙后代”技术的关键构件。

我们还观察到了“雪崩式的替代”。在这里，请回忆一下熊彼特所说的“创造性毁灭的风潮”。这种替代遵循幂次法则，因此技术集合的证据表明它处于自组织临界点上。我们还发现，系统创造出复杂电路的方式是：首先满足最简单的需求，然后将所得结果作为构件，以自展的形式满足更复杂的需求。

电路设计的进化

我们把我们这个人工系统的每一次运行视为一个实验。每个实验都从最原始的组件，通常为一个最基本的逻辑门开始行动，让计算机随机地以非循环的方式将几个组件连接在一起，以便生成新的电路。在运行过程中，所用的组件既可以是原始的逻辑门，也可以是用原始的逻辑门构建出来，并已经封装好的其他电路，即将它们视为具有指定的输入和

输出引脚的芯片。在表7-1中，我们指定了一组“需求”或者说目标，它们是通过组合可以实现的有用的逻辑函数。这些“需求”类似于现实世界中驱动技术进化的真实需求。在理想的情况下，我们希望这些需求是由那些“居住”在这个人工世界中的行为主体产生的。在这个人工世界里，加法器和比较器等被证明是有用的。但是在这里，我们先不考虑这种复杂情况，而是直接列出了一组有用的逻辑函数。当适当的电路出现后，这些逻辑函数就可以实现了。

表7-1　需求表

名称	输入	输出	说明
非	1	1	非
蕴涵	2	1	蕴涵
n 路异或	n	1	异或
n 路或	n	1	对 n 个输入进行析取处理
n 路与	n	1	对 n 个输入进行合取处理
m 路按位异或	$2m$	m	对 n 对输入进行异或处理
m 路按位或	$2m$	m	对 n 对输入进行析取处理
m 路按位与	$2m$	m	对 n 对输入进行合取处理
全加法器	3	2	两位带进位相加
k 位加法器	$2k$	k+1	加
k 位等于	$2k$	1	等于
k 位小于	$2k$	1	比较

注：这是对通用的逻辑函数的一些需求（$2 \le n \le 15$，$1 \le k \le 8$，$2 \le m \le 7$）。

在我们这个人工系统中，既存在对逻辑函数的需求，同时也提供了用来实现它们的手段。采用这样一个人工系统的优点在于，可以容易地对需求和技术进行比对。对特定逻辑函数的每个需求，都可以用特定的真值表来表示：对于每个可能的输入值集合，给出与之相对应的预期输出值的集合。而且，每个创造出来的电路，即每种技术都提供了一个功能，它也可以用真值表来表示：对于提供给输入引脚的每一组二进制值，在输出引脚上产生特定的二进制值。这样一来，我们很容易就可以将实验中的技术与我们的需求表匹配起来。我们完全可以把一种技术的行为、它的真值表视为这种技术的“表型”，而其基因型则是实现这种功能的架构或内部电路。多种不同的基因型可以产生相同的表型。

因此，我们的计算机模型包括了：一组原始的电路、一组利用原始

电路和其他组件构建而成的技术或组件，以及一组有待满足的需求。通常来说，我们只使用一个原始电路，即一个与非门，其表型为¬（$x \wedge y$）。从根本上看，我们的实验是非常简单的。在每个进化步骤中，都会有新的电路从现有电路中创建出来，方法是根据一个用来指定选择概率的选择函数，从所有先前存在的技术中选出2至12个电路来，随机组合布线。同时，通过以不同顺序选择不同内部布线作为输出引脚，创建出新电路的不同表型。在每一步，都存在一组最能满足所有需求或目标的现有技术。在它们的真值表中，不正确元素的数量最少。因此，以这组现有技术为基准，对于每个候选电路，都可以有针对性地进行测试，以便搞清楚它是否对它们有所改进。改进大体上可以分为两类：第一类是这个候选电路与需求的真值表的匹配更好，第二类是它实现了同样的功能，但是比现有的电路成本更低。电路的成本取决于它的部件的数量和各个部件的成本。在上述两种有所改进的情况下，这个电路不仅会把它直接改进的那个电路替代掉，而且也会替换掉所有以那个电路为组件的电路。然后，这个电路将被封装起来，成为一个新的组件，在进一步的组合中，它可以作为一个构件。依此类推，一系列封装好的技术就被构建出来了。如果找到一个完全满足某种需求的真值表的新技术，那么我们就说这种需要得到了“满足”。当然，新创建的电路不能替代自己的某个组件。有用的组件会被赋予一个名字，如“tech-256”或“全加器-121”，并且可以在更高层次的技术中使用。能够完全、准确地满足某种需求的组件，则赋予一个便于记忆的、可以描述该种需求的名称，如三位加法器。用来实现这些一般算法的详细步骤，以及更多的细节，将在下一节中给出。

不过，关于这个人工系统与现实世界之间的对应关系，我们还要多说两句。现实世界中的新技术确实是现有技术的组合，但是现在几乎没有任何新技术是通过随机地将原有组件组合起来而发明出来的。但是，我们不难把我们这个进化过程的每一步，都想象为一系列用来深入探究一个新想法的、严格的实验室测试。或者更确切地说，我们可以认为，我们的过程对应于现代组合化学或合成生物学中所运用的那种过程：先以随机组合的方式产生新的功能，然后再对它们的有用性进行检验。在这个过程中，会出现一个不断增大的有用元素库，这些元素可以用于进一步组合。

更一般地说，我们还可以将这个过程视为一种算法。这种算法不是用来解决特定问题的，而是用来构建一个可以用于进一步组合的、由有用功能构成的“库”。这个算法很好地模拟了现实世界的技术进化过程，

即首先构建能够满足简单需求的技术，再以这些技术为构件，一步步地将更加复杂的技术构建出来。

新技术的涌现

在我们的基本实验设计中，当经过了25万个进化步骤之后，被发明出来的最复杂的电路如下：

> 8路异或、8路与、8路或、3路按位异或、4位等于、3位小于以及4位加法器。

而在一个更加精简的实验设计中，当经过了25万个进化步骤之后，这个人工系统创造了一个8位加法器。它能够正确地进行8位加法，这并不是一个轻而易举的事情。在基本实验设计中，上述电路被发明出来的顺序在不同的实验中是不同的，而且也不是所有这些电路都一定会在一个实验中进化出来。

在实验的初期，一些简单的目标得到了实现。从图7-1中，我们可以看出，即便是对于简单的电路，也出现了一些并非显而易见的实现方法。这些电路随后被封装起来以备进一步地使用。

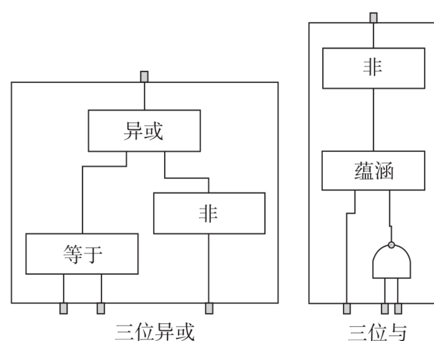


图7-1　为了实现简单目标而"发明"的两种电路

随着进化过程的进一步展开，更复杂的电路开始由更简单的电路中被构建出来。如图7-2所示，二位加法器电路的出现，需要利用支持技术TECH-712。而后一种电路正是"中间技术"的一个很好的例子：它对满足二位加法器这个特定目标有用，但是仅凭自身是不能满足这个目标的，因为低阶位输出，即左侧位输出的计算是不正确的。二位加法器的正确电路由TECH-712电路，再加上一个能够修正该误差的电路构成。

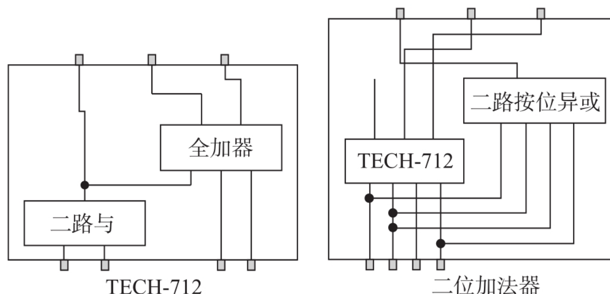


图7-2　利用TECH-712实现“2位加法器”的目标

因为低阶位在左侧，对于多位加法器，各输入位是交叉的。

在我们这些进化出来的电路中，还包括了部分未被使用的电路。图7-3右侧所示的三位加法器就是一个例子。在实验过程中，这种“冗余”通常很快会消失，因为成本“不那么昂贵”的电路，很快就会替换掉“不必要地复杂”的电路。

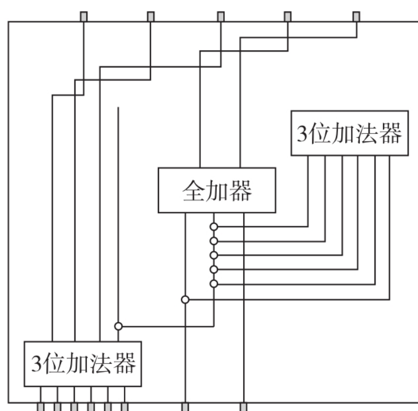


图7-3　带有一个未连接模块的“4位加法器”电路

我们这些实验都是从“与非”这种原始电路开始的。而在另外一些实验中，我们则以“蕴含”电路为原始电路。同样地，在那些实验中，复杂的电路也进化出来了。该过程从“蕴涵”这种原始电路构建出了“非”“与”“异或”，然后进一步构建出了更加复杂的电路。

乍一看来，像8位加法器这样的电路自动涌现出来，似乎也不是什么特别了不起的事情。但是我们一定要考虑到组合的巨大可能性。如果一个组件有n个输入和m个输出，那么就会有（$2m$）$^{(2n)}$个可能的表型，而且每一个表型都可以通过大量不同的电路得以实现。例如，8位加法器具有16个输入和9个输出，它只是有16个输入和9个输出的这类电路超

过$10^{177\,554}$个表型之一。由此可见，通过25万个步骤的随机组合发现8位加法器这种电路的可能性，是完全可以忽略不计的。因此，我们的实验或者说算法发明复杂电路的途径只能是：首先满足更简单的需求，并将其结果作为构件以自展方式去满足更复杂的需求。

技术的扩展

要讨论技术的扩展（build-out），我们需要先给出两个定义。首先，我们把所有用到过的手段和设备，即我们这个实验中的所有电路的集合称为“常设技术储备”（standing reserve）。其次，我们把当前可用的，以及现在正被使用的，且没有被替换掉的技术的集合称为“活跃技术储备”（active repertoire）。

在图7-4给出的三条曲线中，有一条显示的是随着时间的推移，“常设技术储备”，即所有被发明出来的技术的集合的扩增曲线，那是一条单调上升的曲线。相比之下，表示“活跃技术储备”，即当前实际使用的技术的数量的曲线，则不是单调的。这表明重要的发明使得旧的技术过时了。图7-4还表明，系统满足真实功能“需求”的能力得到了持续的改进，这一点从不断增加的替换数可以看得很清楚。

图7-4的时间轴，即横轴上有一些标记，它们标明了一种需求得到了满足的时间。一开始，进展相当缓慢，在实验运行的早期，也完全没有达成任何目标。在一段时间之后，一些功能性“物种”开始出现，并导致更多“物种”的出现。然而，进化过程并不平滑。从图形直观上看，整个过程被不时出现的若干个组分隔开来了。这是因为，每当一项关键技术，即关键的构件或组件被“发现”后，就会快速地“启动”其他技术。例如，在发明了一个“或”电路之后，可以用于3位“或”、4位“或”、5位“或”，以及按位“或”运算的电路就紧随而至了。这些关键构件一旦出现，就能够迅速地激发大量进一步的技术出现。它们在现实世界也有与之相对应的例子，如读者可以想一想蒸汽机、晶体管或者激光，生物进化中物种的扩增也有类似之处。

发明的顺序也很重要。虽然“非”是比“蕴涵”更简单的逻辑函数，但是在某些实验中，“蕴涵”电路可能先被发现，并被作为关键构件使用。

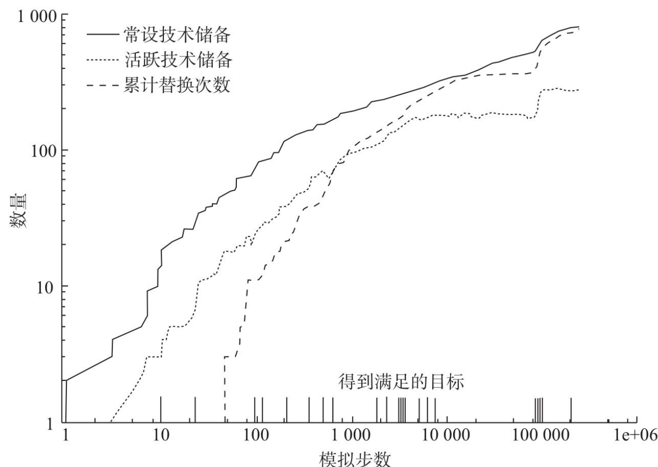


图7-4　“常设技术储备”与“活跃技术储备”的替换次数

“常设技术储备”是单调增长的，而“活跃技术储备”则不然。这是因为，新发明可以改进，并可能取代好几个现有发明。

图7-5显示了一个以“蕴涵”电路为原始电路的实验结果。从图中可见，在开始阶段，“蕴涵”电路被大量用于其他技术，而且比“非”电路更加普遍。但是到了最后，随着成本更低的“非”电路以及其他电路更多地出现在组合中，“蕴涵”电路作为一个组件被使用的频率下降了。为了便于比较，该图还显示了第三种技术，即TECH-69电路。这种技术也要执行“蕴涵”功能，但是还有三个额外的冗余输入，并包含了不需要的组件。最终，所有用途上的TECH-69电路都被功能等效同时又更加有效的“蕴涵”电路取代了。

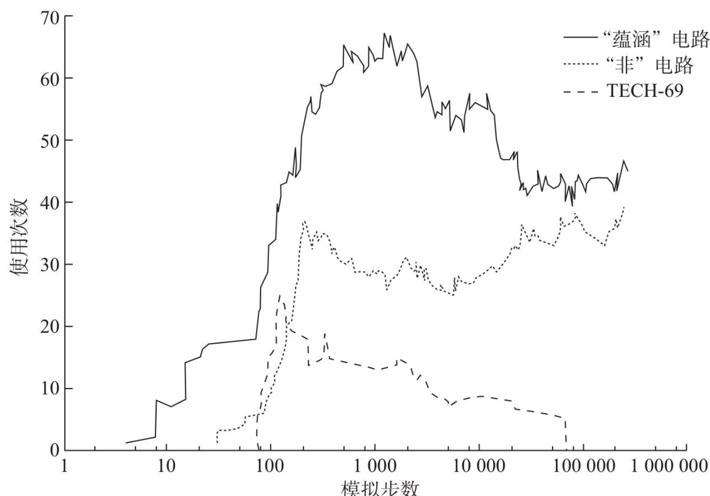


图7-5　“蕴涵”电路与“非”电路被使用的变化图

在本例中，“蕴涵”电路的被发明，早于“非”电路的被发明，而且被使用得也更多。不过，随着时间的推移，当更好的技术出现后，“蕴涵”电路的使用量随之下降。

在公布的需求或目标的数量与新技术的创造之间，存在着一种权衡。为了说明这一点，我们运行了这样一个实验：掩蔽掉一些需求，同时保留一个我们认为对构建加法器有用的子集，如“非”“蕴涵”“二路或”“二路异或”“全加器”，以及k位加法器，其中$1 \le k \le 8$。要想简化进化过程，我们可以这样做：只有当更简单的需求得到满足之后，才加入更加难以满足的需求，难度用输入和输出的数量来度量。结果表明，8位加法器进化得非常快：只要6.4万个模拟步骤，这种加法器就涌现出来了。相比之下，在使用更一般的目标时，有几个实验中需要运行超过67.5万步，才会有一个4位加法器进化出来。大量的不同需求，导致不同功能广泛地生成于电路设计空间当中，但是满足特定复杂需求的速度则会缓慢得多。聚焦的目标会导致更加深入的搜索，因而满足特定的复杂需求的速度也会快得多，但是只能生成更狭隘的功能库。

我们这个算法在不存在中间需求时不会创造出复杂电路。如果我们在一开始时没有设定这类需求，那么必要的构件库就无法生成。例如，如果在上面列出的加法器的目标一览表中略去“全加器”目标，那么实验在运行了100万步后也无法生成二位加法器。而当存在“全加器”这个目标时，就完全不一样了。甚至，二位加法器偶尔会出现在“全加器”被发明之前。这是因为，构建“全加器”所需的技术，对于二位加法器也是有用的。

在实验的每个步骤，只有组合了少于12个现有组件的电路，才会被作为候选电路考虑。这个事实确定了任何一个时刻的可能实验电路集合。这个集合的元素数量当然要比12大得多，我们可以从相邻可能性的角度来分析。我们可以将这个集合视为围绕现有技术的一团“概率云”，随着靠近中间目标的那些点渐次得以实现，这团“概率云”逐步接近于新的技术。因此，如果目标过于复杂，那么就无法以合理的概率达成，或者说实现。同样地，如果不存在作为“垫脚石”的中间目标，那么算法也无法奏效。

更复杂的技术总是用更基本的技术构建的，而且这是一个简单递归重复的模式。在这个意义上，我们列出的作为需求或目标的这些技术，可以说是“有序”的。例如，像4位加法器这样的复杂电路，可以用诸如加法器和半加器这样更简单的元件重复地组合构建而成。

那么，如果我们选定的复杂目标不容易通过上述这种重复的模式实

现，又怎么办呢？为了做到这一点，我们可以选择有n个输入和m个输出的随机真值表，把它们作为需求。在实验中我们发现，这些目标通常不能通过我们的标准中间步骤实现，这个结果不足为奇。那么，如果我们用相同“中间尺度”的随机真值表，替代我们的中间“垫脚石”目标，又会怎样呢？同样地，这么做也不能得到很好的结果。我们这个算法在各种需求被排好序的空间里，即可以通过重复模式来满足最为有效。在这种条件下，复杂性可以充分利用“从更简单的对象构造更复杂的对象”这个规则，以自展的方式成长起来。

创造性毁灭的风潮

当前用于构建进一步的技术的每一种技术，或者说封装电路，都是活跃技术构成的网络中的节点。如果两种或更多种技术，直接被用来创建一种新技术，那么它们就与该技术之间建立了定向链接。因此，给定一种技术A，它必定会链接到它自己的“用户技术”（user technology）上去，用户技术有了这种技术后直接变成可能的那些技术。如图7-6所示，其中一些技术拥有许多链接。这意味着，这些技术大量用于构建新的技术，其他技术则只有很少的链接。链接的分布接近符合幂律，即会生成无标度网络，当然绝不是完美的幂律。

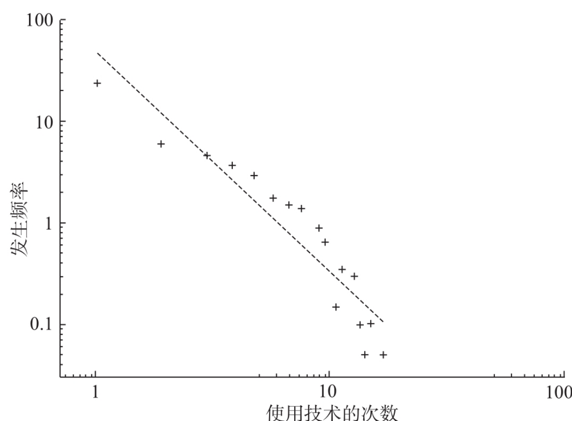


图7-6 极少数关键技术被大量直接用于构建新技术

图中给出了20多个实验的平均结果，每个实验都运行了25万步才结束。

有时，系统会发现某种新的、更优越的执行某个功能（真值表功能）的方式。这就是说，有的新电路能够执行同样的功能却只有更少的组件，或者有同样多的组件却能更好地执行该功能。在这种情况下，新

电路就会将旧电路替换掉。

而且，这种替换将发生在所有“后代”电路中的旧电路上。在网络中，所有直接或间接地将那个旧技术用作组件的技术都要替换掉。在我们的算法中，替换是立即进行的。

当然，技术替换也可能导致网络中的技术出现倒退。假设Tech-124用于构建Tech-136。然后系统发现，要实现Tech-136的功能，有一种更加优越的方法，那么Tech-124就可能被弃置一边不再利用。因为，它既可能无法满足任何目标，也可能无法用于任何活跃技术。在这种情况下，Tech-124就会从活跃技术集合中消失。

随着它的消失，它的一些组件技术也可能不再被使用。于是这些技术也从活跃技术集合中消失。这样一来，大量技术可能会雪崩式地从活跃技术集合中“扫地出门”。这种或大或小的雪崩，就是熊彼特所说的“创造性毁灭的风潮”。

图7-7描绘了这种沙堆崩陷式的技术弃用情景，它遵循幂律。然而，规模轴，即横轴上的那种尺度不会延伸得太远，因为网络中的技术数量从来不会太大。不过，我们还是可以说，图7-7表明，我们这个技术系统存在于自组织临界点上。

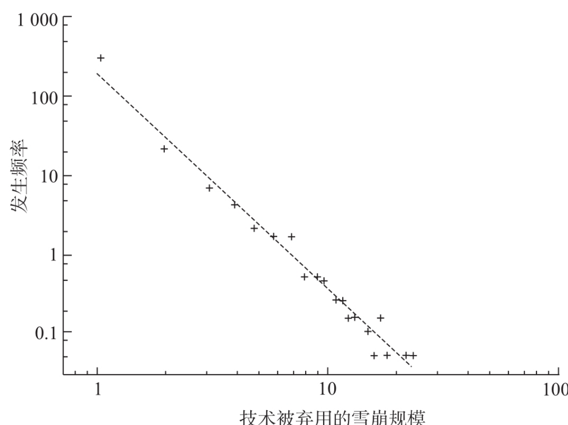


图7-7 创造性毁灭的风潮

图中给出了20多个实验的平均结果。

复杂技术是涌现出来的

利用人工系统，我们已经阐明了技术是如何以自展方式，在自我引

导下从极端简单进化为相当复杂的。这种复杂性，既体现在创造出来的对象的数量上，也体现在这些对象自身的性质上。当然，在现实世界中，新技术通常不是通过随机组合构建出来的，而且也不存在那种指明需要创造技术的公开的需求表。然而，所有新技术确实都是通过对已经存在的组件和部件进行组合而构建出来的。它们所满足的需求，通常无论在经济上还是在技术上，都清楚地发生了信号。并且现有的技术也构成了未来技术的基础，或者说形成了一个构建未来技术所需的构件库。

这个模型有力地刻画了我们在现实生活中观察到的某些现象。大多数技术都不是特别有用的构件，但是某些使能技术，如激光或晶体管，是创建“后代技术”的关键。在我们这个模型中，我们发现有强烈的证据表明，我们的活跃技术集合在统计特征上与地震或沙堆相似，这也就是说，它存在于自组织临界点上。我们的模型还表明，技术的扩展在很大程度上取决于为中间需求或更简单的需求而构建出来的早期技术是否存在。这也印证了伦斯基等人的发现：在生物系统中，复杂的特征是可以涌现出来的，但是只有在更简单的功能已经得到了进化女神的青睐并发挥着“垫脚石”作用的条件下。

COMPLEXITY AND THE ECONOMY

08

技术进化所引发的经济进化

经济是从它的技术中涌现出来的

众多的技术集合在一起，创造了一种我们称之为“经济”的东西。经济从它的技术中浮现，不断地从它的技术中创造自己，并决定哪种新技术将会进入其中。每一个以新技术形式体现的解决方案，都会带来新的问题，这些问题又迫切需要进一步得到解决。经济是技术的一种表达，并随这些技术的进化而进化。

新技术进入一个经济体后，不仅会使经济增长，而且会使经济变得更有生产力。如果对经济的影响很重要，那么新技术还会导致经济在结构上发生变化，并且经济得以组织起来的方式也会变得有所不同。那么，这种结构变化是怎样发生的？经济又是如何随着技术的进化而进化的？

本章认为，新技术的出现会触发一系列级联式的事件。新技术会成

为进一步的新技术的构件。新技术会创造出对支持技术、新型组织形式以及新型机构的需求，从而创造出机会利基。新技术也可能带来社会和经济挑战，从而需要进一步的技术来解决这些挑战。因此，新技术带来了一连串的进一步的技术，经济的特性或内部结构也将因此而不断变化。本章原是2009年出版的著作《技术的本质》中第10章“技术进化所引发的经济进化”的内容，在此略有改动。

* * *

在《技术的本质》一书的第10章中，我直接以技术的进化为主题进行了讨论。不过，还有另一种感知技术进化的方式，那就是通过经济的“眼睛”去观察技术的进化。经济能够映射出我前面一直在讨论的技术的变化、增长和替换。经济对技术的这种映射，正如在《技术的本质》第8章中所说的，不仅是通过平滑地调整生产模式和消费模式，也是通过创建新组合实现的。经济会随着技术的进化而改变它自己的结构，即改变它的制度安排方式。这种改变时时刻刻都在发生，而且发生在所有层次上。

因此，我想从经济的角度出发，重新回过头去分析技术进化的各个步骤，看看它们是如何依次展开的。我们将会看到经济的结构性变化的自然过程。这个过程是由我们刚刚深入探究过的技术进化过程驱动的。为了阐述技术进化驱动了经济进化，我们首先必须做到的一点是：以一种不同于标准经济学的方式来思考经济。

经济就是技术的一种表达

无论是在词典中，还是在经济学教科书中，定义经济的标准方式就是将它定义为一个“生产、分配和消费”商品和服务的“系统”。[1]这种定义将“经济”这个系统当成了一个背景：经济本身之所以存在，就是为了向发生在其中的事件和调整提供背景。这样一来，经济就变成了一个巨大的容器，里面装载着它的技术，或者说，是一台包含着许多模块和零部件的巨大机器。这些模块和零部件就是经济的技术，或者是它的生产手段。当出现了一种新技术（如铁路交通）后，就为某个特定行业提供了一个新的模块，并使这个行业得到升级：新技术替换下来的旧的专用模块（在铁路交通的兴起这个例子中是运河）将会被舍弃，而新的升级模块则悄然到位。至于这台“机器”的其他部分，则会自动地重归平衡，

而且它的张力和流量，即价格及生产和消费的商品和服务，都会相应地进行重新调整。

这种看法不能说全错。而且，我在大学里读研究生时，被灌输的就是这种思考经济学的方式。这也正是今天的经济学教科书描述经济时所用的方式。但是，这种看法也不能说全对。为了探索结构变化，我将以一种不同的方式来看待经济。

我将把经济定义为一整套安排和活动，社会需要借助这一整套安排和活动来满足自己的需求。有了这一定义，经济学将研究这一整套安排和活动。那么关键是，这些安排到底是什么？对此，我们可以从维多利亚时代经济学家所说的“生产手段”，即处于经济核心的工业生产过程开始讨论。我的定义应该不会令卡尔·马克思觉得惊讶。很早之前，马克思就指出经济形成于它的“生产工具”，而他所指的生产工具，则包括了他那个时代的大型工厂和纺织机械。

但是，我将不限于马克思所强调的工厂和机械。形成经济的一整套安排，包括了数之不尽的所有设备和手段，还包括了被我们称为技术的所有目的性系统。这就是说，包括了医院和外科手术、市场和定价系统、贸易安排、分配制度、组织和企业，还包括了金融体系、银行、监管系统和法律制度。所有这些，全都是我们用来满足我们的需求的安排；所有这些，都是实现人类目的的手段。在《技术的本质》开始的章节中，我把技术定义为实现人类目的的手段，因此所有这些都是“技术”或目的性系统。由此可见，如果我们承认纽约股票交易所和合同法的专门规定，与钢铁厂和纺织机械一样，都是实现人类目的的手段，那么我们就可以说，它们在广义上也都是技术。

如果我们把所有这些“安排”都纳入技术的集合当中，我们就会看到，经济不是技术的容器，而是建构在技术之上的。经济是一系列活动、行为、商品和服务流，是以技术为中介或被技术所覆盖的。也就是说，我一直在讨论的那些方法、过程和组织形式构成了经济。

经济是其技术的表达。

我并不是说一个经济体与它的技术是完全相同的。经济所包含的东西还有很多。商业中的谋划、投资、投标及交易，这些都是活动，但并不是目的性系统。我要说的是，经济的结构是由它的技术形塑的，或者如果你喜欢的话，也可以说技术构成了经济的骨架。经济中的其他部分，如商业活动、博弈中各方参与者的策略和决策、商品和服务的流

动，以及随之而来的投资等，则构成了经济体的肌肉、神经和血液。但是，所有这些“其他部分”都只是环绕在经济的外围部分，而且它们也是由技术，即目的性系统形塑而成的。只有技术才构成了经济的结构。

我在这里建议的这种思维转变的幅度其实并不大，但是确实很微妙。这就像我们不再把心智视为装载观念和惯常的思维过程的容器，而要把它视为从观念和惯常的思维过程涌现出来的东西。或者说，这就像不再把生态环境视为容纳各种生物物种的容器，而是把它视为物种的集合的产物。经济也是如此。经济因它的技术而形成了一种生态。经济形成于技术，这就意味着经济不是单独存在的。就像生态为物种形成了生态位一样，经济为新技术形成了机会利基，当新技术出现后，这些机会利基就会被填补上。

思维方式转变了，结果也就不同了。因为这意味着，经济是从它的技术中涌现出来的，或者说萌发出来的。这还意味着，经济不仅必须随着技术的进化而重新调整适应，而且还必定会随着技术的进化不断地形成和重构。这也意味着，经济的特征，即它的形式和结构必定会随着技术的变化而变化。

总结一下，我们可以这样说：当技术构成了集合之后，创造出了一个结构，决策、活动、商品流和服务流全都发生在这个结构中，进而创造出了被我们称为“经济”的那种东西。因而，经济就是以这种方式，从它的技术中涌现出来的。它不断地从它自己的技术中创造自己，并决定哪些新技术将进入它自身。请读者注意，这里的因果循环是这样的：技术创造了经济的结构，同时，经济调节着新技术的创造，因此也调节着经济自身的创造。我们通常看不到这种因果循环：



这是因为在短期内，如一年左右的时间，经济似乎是固定不变的。它确实很像它的活动的一个容器。只有从几十年的长期的角度观察，我们才能看清楚那些使得经济形成、互动，以及再次崩溃的安排和过程。只有在更长的时间跨度内，我们才有可能观察到一个持续创造和再创造的经济。

结构性变化

当新技术进入经济后，这个从它的技术中形成的系统会发生什么呢？当然，我们仍然会观察到，我在《技术的本质》第8章中谈到过的那种调整和新组合。它们仍然完全有效，都会对经济产生很大影响。但是除此之外，我们还可以看到更多的东西：增加新技术，会推动经济结构发生一系列变化，也会给经济的制度安排带来一系列变化。

现在，我们进入了一个经济学理论通常不会进入的领域：结构性变化。但是，这个领域并不是一个完全空白的“无人地带”。事实上，历史学家们，更准确地说是经济史学家们，一直在这个领域“精耕细作”。历史学家们在看待经济时，向来都认为新技术进入经济，不仅会引发经济的重新调整适应和增长，而且会导致经济的构成，即经济的结构发生变化。但是，大多数历史学家们进行的都是案例研究，得出的是只适用于特定的情况的特殊理论。相比之下，我们对经济和技术的思考方式，却给我们提供了一个抽象地思考结构变化的途径。

在现实世界中，一项新技术很可能会创造出一个新的行业。它可能会要求建立新的制度安排，也可能会引发新的技术和社会问题，并且因而创造出新的机会利基。而且，所有这些反过来又可能引发进一步的组合变化，即新技术。如果我们借用《技术的本质》第9章中所总结的技术进化中的步骤，并通过经济的视角来重新审视这些步骤，那么我们就可以刻画这种变化序列。接下来，我们就来这样做。

我们可以从假设一项“候选技术”出现在了经济中开始。这项技术是通过对以前的技术的组合创造出来的，而且已经击败了它最强劲的竞争对手。现在，它进入了经济，接下来的事情可以总结为6个事件，或者说步骤。我们可以把这些步骤看作这个“技术建构博弈”的“法定行动”。接下来，我将在抽象的层面上进行阐述。但是读者可能会发现，如果在思考时，心中想象着某种可以作为案例的技术，如铁路、汽车或晶体管，那么将会对理解这些步骤大有帮助。

1．新技术作为一个新的元素进入活跃技术集合，成为了活跃技术集合中的一个新元素。

2．该新元素可用于替换现有技术或现有技术的组件。

3．该新元素为支持技术和制度安排，建立了进一步的“需求”或机会利基。特别是，这种需求可能源于新技术导致的新的技术、经济和社会问题。因此，新技术带来了对能够解决这些问题的进一步的技术需求。

4．如果旧的、被替换掉的技术从技术集合中逐渐“隐退”，那么它们的随附需求也会逐渐消失，当然它们提供的机会利基也将随之消失，因此填补这种机会利基的那些元素从此将变得不再活跃。

5．新元素作为未来进一步的技术，即未来元素的潜在组件，将变得活跃起来。这也就意味着，新元素会“调动”使用它和容纳它的其他技术。特别是，它可能导致包含它的新组织的产生。

6．经济，即商品和服务的生产和消费模式，重新调整以适应这些步骤。成本和价格以及随之而来的对新技术的激励，也会相应地发生变化。

以上是我用经济学术语，阐述了我们在第9章中描述过的各个步骤。技术进化的机制就是通过这些步骤形成的。但是我现在的目的是，重新解释这些步骤，来描述经济的新结构是如何形成的。当一项新技术进入经济后，它就会召唤新的安排，即新技术和新的组织形式。新的技术或新的安排，反过来又可能引发一系列新问题。为了找到这些新问题的答案，又需要进一步的新安排，或通过修正现有技术来实现这个目的，而这就又打开了对更进一步的新技术的需求之门。所有这些，构成了一个不断向前推进的“问题与解决方案（挑战与回应）”的序列，而这也就是我们所称的结构性变化的序列。就是通过这样一种方式，经济在暴风骤雨般的变化过程中构成了自身并不断重构自身，这包括为了适应这种变化的创新和新的制度安排，以及因此而打开的、相互之间如影随形的机会利基。

下面，让我用一个具体的例子来详细说明这一点。在18世纪60年代的英国，实用的、易操作的纺织机械出现了，这就提供了一种可以替代当时以手工作坊为基础的生产方式的途径。当时流行的生产方式是散工制或外放分工制（putting-out system）：羊毛和棉花的纺织都是在家里以手工方式完成。[2]但是，新的纺织机械在刚开始时只取得了部分成功，因为它要求比家庭手工作坊规模更大的生产组织。这样一来，纺织机械就为更高级的组织安排，即纺织厂提供了机会，并使自己也成了更高级组织中的一个组成部分。工厂本身作为一种组织手段，这也是一种技术，反过来也要求另一种手段来实现机器生产，那就是工厂劳动力。劳动力当然早就已经存在于经济当中了，但是当时的劳动力数量并不足

以支撑这种新的工厂生产体系。所需要的劳动力主要是从农业转移出来的，这种转移又反过来要求在工厂附近有居住的地方。因此，工人宿舍和工人住房的建设就开展起来了。而工厂、工人及工人住房的组合，就是工业城市。随着工业城市的增长，一整套新的社会组织方式，或者说一整套新的制度安排出现了。由此，维多利亚式的工业经济的结构就开始涌现出来。是的，一个时代的特征，即一整套与机械化工业生产的优越技术相匹配的制度安排，就是以这种方式逐渐呈现出来了。

但是，任何一个时代都不会就此终结。在当时的制造业工人中，有许多是童工，他们经常在狄更斯作品中所描述的那种条件下工作。这种情况引发了强烈的改革需求。人们不仅呼吁提升“下层阶级的道德水准”，而且要求保障他们的安全。[3]于是，到了时机成熟时，法律制度做出了进一步的调整，以回应这种呼声。例如，旨在防止最糟糕的情况发生的劳动法出台了。然后，新的工人阶级开始要求在工厂创造的财富中分享到更大的份额。为了改善自己的境况，他们找到了一种有效的手段，那就是组织工会。工厂中的工人比分散在孤立的家庭小作坊中的劳动者更加容易组织起来，因此在短短的几十年内，工会就成了一支强大的政治力量。

通过这种方式，纺织机械的到来，不仅替换了家庭小作坊的手工生产方式，而且还为更高一级的制度安排创造了机会，那就是工厂生产体系。在这个制度安排中，机械只是其中的一个组成部分。新的工厂系统反过来又生成了一个对劳动和住房需求的链条。而为了满足这些需求所提出的解决方案，又创造了进一步的需求。所有这些，到最后进化成了维多利亚式工业体系。整个过程用了100多年的时间才最终完成。

有的读者可能提出反对意见说，这种描述使得结构变化看起来太简单化了，或者说太机械化了。技术*A*提出了一个对安排*B*的需求，技术*C*实现了这个安排*B*，然后技术*C*又提出了进一步的需求*D*和*E*，而这些需求则由技术*F*和技术*G*来满足……当然，确实是这样一种序列构成了结构性变化的基础，但是这样一种序列决不是简单的。关于这一点，只要我们思考一下，这些新的安排和技术本身所引发的次一级技术和次一级安排的递归循环过程，就可以明白了。工厂体系本身就需要为新机器提供动力的手段，需要传递动力的绳索和滑轮系统，需要获取和跟踪材料的方法、新的记账方法、新的管理手段、新的产品交付方式，等等。所有这些又都是用其他组件构建出来的，因而都会产生自己的需求。结构性变化本身是一种分形，它会在次一级的层级上伸展出进一步的“分支”，

就像胚胎动脉系统会随着胚胎的发育，而分支出更小的动脉及毛细血管一样。

技术进化引发的反应当中，有些根本不属于经济领域。手工作业可以机械化这个观念，很快就从纺织业传播到了其他行业，并导致了新的机械设备的出现。从心理学的角度来看，工厂体系不仅创造了一种新的制度安排，而且要求一种“新人”的出现。历史学家戴维·兰德斯指出：“工厂纪律需要并最终创造了一种新的工人……纺纱工不能再在自己家里自在地摇她的纺车，织布工也不能再在自己家里轻快地抛他的梭子，因为不受他人监督的美好时光已经一去不复返了。现在，工作必须在工厂中完成，工人必须跟上由那些永不疲倦但了无生气的机器设备规定的步调。工人成了一个大团队中的一员，这个团队必须共同进退，开始、歇息和结束，一切都在监督者的密切注视之下进行。为了让工人勤奋工作，工厂主采取了道德劝勉、金钱奖罚等手段，有时甚至不惜通过伤害工人的肉体去强迫他们。工厂成了一种新的监狱，而时钟则成了一类新的狱卒。”因此，新技术不仅仅导致了经济变化，它还造成了心理变化。

在讨论经济的结构性变化时，我们还必须记住一点，这就是变化不一定全部都是有形的，也不一定全部都是某种“安排”。我们还需要记住，变化可能是多因多果的。不过，我最想强调的是，对于经济的结构性变化的过程，只要利用上面列出的描述技术进化的各个步骤，我们是可以从逻辑上，或者说，从理论上进行探究。经济的结构性变化不仅仅是新技术的增加和旧技术的替代，以及紧随其后的经济进行重新调整适应。相反，它是一连串后果结成的链条。在这个链条中，构成了经济结构的骨架的那些安排不断地召唤新的安排。

当然，对于那些制度安排，它们要落到实处并确定经济的结构，没有什么是不可避免的，或者说没有什么是预先决定的。我们早就知道了，有很多不同的组合，有很多种制度安排，它们都可以解决技术带来的问题。到底哪一个被选中，则取决于历史上发生的小事件，如问题出现的顺序、个人性质倾向和行为偏好等。技术决定了经济的结构，并进而决定了由此产生的大部分世界，但是到底哪些技术能够“落地”，却不是事先就能确定的。

解决方案带来的新问题

我在前面已经指出过，结构的展开也是构成经济的那些安排不断重构的过程。一套安排为下一套安排的到来创造了条件。这种重构过程一旦启动，就没有结束的理由。即便只有一项新技术，读者不妨想一想计算机或者蒸汽机，这种后果序列也很可能会连绵不断地延续下去。

反过来说，经济也永远不会停滞。在任何时刻，经济的结构都可能处于高度相互兼容的状态，因此看上去接近不变。但是事实上，在这种静态的表面之下，早就蕴藏了“自我毁灭”的种子，就像熊彼特在100多年前指出的那样。原因是创新。一旦经济创造出了新的组合、新的安排或者熊彼特所说的“新的商品、新的生产或运输方法、新的市场、新的工业组织形式”，“工业突变”过程就会出现。这种突变“不断地从内部彻底地重构经济结构，不断地破旧立新”。

是的，系统总是“时刻准备着”从内部发生变化。

不过，与熊彼特相比，我要在这里阐述的观点有更多的含义，而且不仅仅只是多一点。新技术的到来通过发现新组合打破了现状，这些组合是比我们现在使用的商品和方法更好的“新版本”。不仅如此，为了容纳新技术，需要进行一系列的调整适应，从而引发一连串的新问题，这样也就创造了新的机会利基。而新的机会利基又会召唤更进一步的新组合，进而引入更进一步的新技术和更进一步的新问题。[4]

因此，经济总是存在于永恒的变化之中，总是存在于永恒的新奇之中。这也就是说，它永远存在于自我创造的过程之中。经济永远不会“满意”。事实上，我们还可以再加上一句：重要程度各异的大量新技术随时随地都会进入经济，不仅如熊彼特所说的“打破了均衡”，而且还会引发连锁性的不断变化。所有的变化都重叠在一起并相互作用，从而触发进一步的变化。因此，其结果是变化带来变化。

然而奇怪的是，对于这种随时都在发生的不断变化，我们却不一定很在意。这是因为，结构性变化的过程动辄长达几十年，一般不可能在短短几个月内结束。它更像在我们脚下发生的缓慢的地质运动。在短期内，现有结构具有高度的连续性。它是一个松散的、有很强兼容性的系统，行为主体在这个系统中通常可以从容制订计划并采取行动。但是在任何一个时刻，这个结构都在改变，经济永远都在建构自身。

那么，这种技术不断进化、经济不断重塑的过程，是不是总有一天会宣告终止呢？从原则上说，这是可能的。但这只是在原则上。只有在未来不能再发现新现象，因而无从创造新技术的情况下，只有在进一步

组合的可能性完全耗尽的情况下，只有在我们人类的实际需求仅用我们拥有的可用技术就可以完全得到满足的情况下，经济才会停步不前。但是，这些“如果”没有一个是可能变成现实的。永无止境的需求和总有可能被发现的新现象，足以推动技术永远向前发展，而经济也将如影随形。

技术进化和经济进化的过程永远不可能停止，还有另一个原因。正如我一直强调的，以新技术的形式出现的每个解决方案，都会创造出新的挑战、新的问题。这可以说是一个一般规律：每项技术都包含着问题的种子，而且通常有好多颗。当然，它不是技术或经济的“法则”，更不是宇宙的“普遍规律”。它只是一个基于人类历史的、有一定普遍意义的经验观察结果，但它还是一个会令人觉得遗憾的观察结果。使用碳基燃料的技术带来了全球变暖问题。使用原子能，即一种环保的清洁能源的技术，则带来了很难处理的核废物问题。而航空交通则意味着病毒在全球迅速蔓延的一种可能性。在经济中，解决方案导致问题，而问题又指向进一步的解决方案。这种在解决方案与问题之间的舞蹈将会一直跳下去，未来的任何时候都不可能改变。如果我们足够幸运，那么我们可以收获净利益，这时我们称之为“进步”。无论是否有进步，这场舞蹈已经注定了，技术将永恒变化，而经济则如影随形。

我在本章中讨论的实际上是从经济的角度看到的技术的进化。因为经济是它的技术的表达，所以经济是一整套安排，这些安排源于各种过程、组织、设备和制度规定，它们构成了不断进化的技术集合，因而经济要随着它的技术的进化而进化。同时，因为经济源自它的技术，所以经济继承了技术的自我创造性、永恒开放性和永恒新异性。因此，经济最终产生于创造技术的现象。这也就是说，归根结底，被组织起来为我们的需求服务的是大自然。

这样一种经济从来不是简单的。任何一个制度安排都是建立在另一个安排的基础之上的：法律制度中涉及商业的那部分，是建立在假设市场和合同存在的基础之上的，市场和合同则必须假设存在银行和投资机制。因此，经济从来不是一个均质的东西。它是一种结构，一个宏伟的结构，由相互作用的、相互支持的安排构成，这些安排存在于各个层次上，很多个世纪以来一直在从自身中发展自己。它可以说是一个有生命的东西，或者至少是一个不断进化的事物。它不断改变自己的结构，因为它的安排会创造出进一步的可能性带来的问题，这就需要进一步做出回应，然后又是进一步的安排。

这种结构的进化是对构成经济的安排的不断重构，因为一套安排为下一套安排的到达创造了条件。这与在给定的安排中或给定的行业内进行调整是不同的，而且也不等同于经济增长。它是连续的、分形的和不可阻挡的。它带来了永不停息的变化。

那么，在这种永不停息的结构性变化中，有什么是不变的吗？有的。经济在形成它自己的模式时，总是要利用一些相同的元素：人类行为的倾向、解释所依据的基本现实，以及购买的商品必定等于销售的商品这样的自明之理等。这些基本的“法则”总是保持不变。但是，它们用以表达自身的手段却会随着时间的推移而变化，并且它们形成的模式也会随着时间的推移而改变和重构。这样，每一个新模式，每一套新安排，都会产生一个新的经济，而旧的结构则如烟般消逝。但是无论如何，经济得以形成的基本组成部分，即那些基本的“法则”却始终保持不变。

经济学作为一门学科，经常受到人们的批评，因为与物理学或化学这样的“硬科学”不同，经济学无法始终“钉”在一套不随时间的推移而改变的描述上。但是，这并不是经济学的失败之处。恰恰相反，这种情况的出现，其实是正当的和自然的。经济不是一个简单的系统，它是一个不断进化的、复杂的系统，它形成的结构永远在随时间推移而变化。这就意味着，我们对经济的解释也一定要随时间推移而不断变化。有时，我会把经济想象为第一次世界大战期间的夜晚战场。那里漆黑一片，当士兵往胸墙外看时，几乎什么也看不到。距我军陆地一公里左右，就是敌军的阵地，从那里传来的声音隐约可闻，可以感觉到敌军正在调兵遣将，进行重新部署。但是我们对敌军的新部署的最佳猜测，也只能是在以往观察到的原部署基础上外推。突然，有人发射了一颗照明弹，耀眼的光芒点亮了整个夜空，整个战场尽收眼底：部队的位置、火力的配备、工事的安排……无一例外。然后，亮光骤灭，一切再次陷入黑暗。经济也是如此。在经济学的历史上，也曾经闪现过一道道耀眼的光芒，那是一些伟大的经济学理论，由斯密、李嘉图、马克思和凯恩斯，还有熊彼特等伟大的经济学家提出。这些理论的思想光芒偶尔会将“经济学战场”照亮一段时间，但是黑暗始终笼罩着。而在黑暗之中，骚动和重新部署一直在继续。我们确实可以观察经济，但是我们所用的语言、我们给它加的标签、我们对它的理解，却全都凝固在了夜空被照亮的那一瞬间，尤其是最近一次，当一组照明弹同时闪耀的时候。

COMPLEXITY

AND THE

ECONOMY

09

复杂性的进化

是越来越复杂，还是随时可能坍塌

我们经常认为，随时间进化的系统，一般都会变得越来越复杂。实际上却并非如此。复杂性随系统进化而增加的机制有三种：协同进化多样性的增加、结构深化和捕获软件。在这三种机制下，复杂性的增加都是间断性的和世代性的。由于前两种机制是可逆的，因此复杂性也可能随时坍塌。

我们经常认为，随时间进化的系统一般都会变得越来越复杂。这似乎是理所当然的、无须解释的。其实不然。1994年，在我写作本章时，对于什么机制可能导致进化有利于复杂性增加，我们的理解相当有限。本章指出，复杂性随着系统的进化而增加的途径有三种。第一种途径是，在共生，即进化的系统中，复杂性可能会通过“物种”多样性的增加而增加：在某些情况下，新物种可以提供进一步的生态位，从而使得更

多的新物种得以涌现出来，导致物种总数稳定地以螺旋向上的形态增加。第二种途径是，在单系统中，复杂性可以通过结构复杂性的增加而增加：系统的内部子系统或子功能或子组分的数量不断地稳定增加，以便突破系统性能限制，或者增强其运行范围，或者用于处理异常情况。第三种途径是，复杂性可能通过“捕获软件”机制而突然增加：系统捕获更简单的元素，并学会将它们编写为“软件”，以便用它们来实现自身的目的。

在这三种机制下，复杂性的增加都是间断性的和世代性的。前两种机制是可逆的，复杂性坍塌的情况随时都可能随机地发生。这方面的例子不仅在生物学中有，在经济学、自适应计算科学、人工生命和进化博弈理论中也比比皆是。本章最早发表在由G.考温（G. Cowan）、戴维·派因斯和D.梅尔策（D. Meltzer）主编的《复杂性：隐喻、模型和现实》（*Complexity: Metaphors, Models, and Reality*）一书中。该书于1994年出版。

* * *

随时间进化的系统一般都会变得越来越复杂，这是一个人们普遍接受的观念，几乎可以说它是一个“民间定理”。但是，这种观点有什么证据吗？进化真的有利于复杂性的增加吗？如果是，那又是为什么呢？进化使得复杂性随时间流逝而增加，是通过什么机制实现的？而且，这个过程会不会向另一个方向发展，即复杂性会不会减少呢？在本章中，我将讨论这些问题。特别是，我将阐明在一般的系统中，进化使得复杂性增加的三种途径。

在生物学文献中，关于进化与复杂性之间的关系，向来是一个聚讼纷纭的主题。但是，大多数争论都流于粗疏，进一步的深入探讨由于如下事实的存在而滞碍难行：进化创新通常以平滑的变化或连续的塑性修饰的形式出现，如体现在生物体的大小、身体器官的形态、生物的行为等方面，因而使得“复杂性增加”很难定义和识别。因此，虽然大多数生物学家都认为复杂性确实会随着进化而增加，而且他们也提出了一些特定的机制，但是由于定义和观察的困难，这个问题仍然深陷在泥沼之中。有一些生物学家甚至怀疑，进化与复杂性之间根本不存在任何联系。

不过幸运的是，近年来，我们积累了不少关于不同进化情境的经验证据，而且它们不全是生物学的。这些情境包括：经济中不同技术之

间、不同厂商之间的竞争，自我复制的计算机程序、自适应计算、人工生命系统，以及由竞争性的博弈策略组成的基于计算机的“生态”。它们不但可以用来替代生物学中的相关实例，而且还拥有两大优点：第一，它们的变化和创新通常是离散的和明确的，所以在这些情境下，我们可以更容易地定义和观察复杂性的增加；第二，它们大多数是基于计算机的，因此它们可以成为“实验室”，使得我们能够实时测量复杂性在进化过程中的变化并进行复制。

本章在讨论复杂性与演变之间的关系时，采用的例子有来自经济学的，也有来自上面提到过的这些情境的，当然还有来自生物学的。我将只关注看上去就是“复杂性”的那种复杂性。“复杂性”完整的准确定义取决于不同的具体情境，但是我希望，随着对复杂性因进化而增加的各种机制的讨论的深入，它的定义将会越来越清晰。我将经常在系统发育的意义上使用“进化”这个术语，因为在一个具有明显的遗传结构谱系的系统中，“发育”可能是随时间推移而展开的。这样一来，我们就可以讨论一种语言或一种技术的进化，而不必先假定该语言或该技术必定会在一个语言“种群”或技术“种群”中得到繁殖。

机制1：协同进化多样性的增加

复杂性随进化而增加的第一个机制，我称之为“共生多样性的增加”（growth in coevolutionary diversity）。这种机制适用于各种个体或实体、物种、生物体在相互作用的种群中共存的系统。在这种共生系统中，一些个体或实体、物种、生物体构成了“基质”，或创造了“生态位”，从而允许其他个体或实体、物种、生物体存在。因此，对于这样的共生进化的系统，我们可以这样来看：它们组织成了松散的层级结构或相互依赖的“食物网”，下一个层级的个体存在取决于更接近“层级基座”的“更基础”的个体存在。

在这些系统中，当个体及它们在互动中的多种可能性创造了大量不会封闭、可以被新产生的个体利用的“生态位”或利基时，多样性倾向于以自我强化的方式增大。而进入种群的新个体又可能会提供新的基底、新的生态位，从而又为更进一步的新个体提供了填补或利用的新的可能性。当然，这些最新进入的个体，也可以提供更新的生态位和基底……依此类推。正是通过这种方式，复杂性随着时间推移而自展地不断增加。在这里，复杂性以更大程度的多样性、更复杂的交互网络的形式出

现。在最初的时候，当新出现的个体仅限于取代没有竞争力的先前存在的个体时，共生进化多样性的增加可能会比较慢，甚至可能会完全止步不前。但是随着时间的推移，当个体开始提供生态位，且生态位使得新个体的出现成为可能时，共生进化多样性就会进入正反馈通道，即多样性本身就可能为进一步的多样性提供“燃料”。

在过去的20年里，计算机工业中的专用性产品和软件出现了快速增长。这种增长就是经济中共生进化多样性增加的一个很好的例子。在现代微处理器出现后，它们就为诸如存储器系统、屏幕监视器和总线接口之类的设备创造了适当的生态位或利基。将它们与微处理器连接在一起，就是非常有用的硬件——计算机设备。而这些设备反过来又为新的操作系统软件、新的编程语言以及新的应用程序创造了需求或利基。这些硬件和软件的存在又使得桌面出版、计算机辅助设计和制造、电子邮件、共享计算机网络等技术成为可能。这些技术又为激光打印机、工程设计软件和硬件、网络服务器、调制解调器和传输系统创造了利基。这些新设备反过来又“召唤”新的更强大的微处理器和系统软件来驱动它们。因此，在这大约20年的时间中，计算机行业经历了多样性的爆炸性增长：从少量的设备和软件到各种各样的设备和软件，因为新设备使得更多的新设备成为可能，新的软件产品使得计算机的新功能成为可能，而这些反过来又要求进一步的新设备和新软件的出现。

当然，我们不应该忘记，随着新的计算机产品和功能的出现，它们也会取代经济中的其他东西。计算机辅助设计可能最终取代基于标准绘图板和T字尺的设计。因此，系统中一部分组件的多样性增加，可能被其他地方的多样性减少所抵消。偶尔在共生系统中，现有功能被替代也可能导致共生多样性增长出现逆转。当新实体替换掉了系统中更基础的实体，同时依赖于被替换掉的实体的利基消失时，就可能发生这种情况。例如，在整个19个世纪，马车运输业出现了许多“利基企业”，它们专业化程度高且相互联系紧密，数量也一直在稳步增加。到了19世纪末，各种不同类型的马车车厢制造者、马具制造商、铁匠铺和马种繁殖场等形成了一个共生进化的生态。不久之后汽车出现了，所有这一切都崩溃了。与之相反，以燃油汽车技术、石油勘探和精炼技术和内燃动力技术等为基础的一系列相互连接的新“利基制造商”，则组成了一个新的网络。这个网络一开始增长缓慢，但它逐渐替代了以往的“马车制造网络”。因此，共生系统中的复杂性，在这种情况下，复杂性就是指多样性，当然可能随时间推移而增加，但是也可能出现很大的震荡。

多样性的增长在很多人工进化系统中都可以观察到。这样的不完全例子包括：汤姆·雷（Tom Ray）开发的Tierra系统、约翰·霍兰德构建的ECHO系统、斯图尔特·考夫曼设计的多个化学进化系统。在这里，作为例子，我简单介绍一下汤姆·雷的Tierra系统。它是雷构建的一个人工世界：不同的计算机程序在虚拟计算机中竞争处理器时间和存储器空间。这个世界是从一个单一的“有机体”开始的，该“有机体”是一组能够自我复制的机器语言指令，偶尔会发生变异。这个“有机体”构成了一个生态位，或者说基底，从而为那些利用它的部分代码进行复制，或者说以它的指令为“食物”的“寄生有机体”的出现提供了条件。此外，还会出现另外一些“寄生有机体”免疫的有机体。“寄生有机体”也会形成基底，供依赖它们的“超寄生有机体”所用。然后，“超超寄生有机体”也会出现……依此类推。就这样，新的“有机体”不断出现、不断消失，形成了一个丰富多彩的、由相互竞争的机器语言程序构成的生态系统。这个生态系统呈现出多样性持续净增长的趋势。在这个系统开始运行后的连续几天里，汤姆·雷一直没有观察到多样性增长的终点。在这个不断进化的基于计算机的系统中，从单个基因型开始，进化出了超过2.9万种不同的、能够自我复制的基因型，它们大致可以分为300个类型，这相当于这个系统中的物种数量。

行文至此，我想向读者指出几个适用于这种机制的要点。

第一，在某些情况下，新实体的出现可能不太依赖于先前实体的存在，而更多依赖于它们之间展开“互动”的可能性。例如，在经济中，只有当激光技术、静电复印技术和计算机技术等技术都存在时，像激光打印机这样的新技术才有可能变成现实。在这些情况下，实体的共生集群，即一组实体，它们的“集体活动”或共同存在有重要意义，可能会创造出很多利基。因此，我们可以预测，当“集体存在”对于利基的形成非常重要时，共生进化多样性的增长在一开始时将会显得非常缓慢。因为在只有极少数实体的情况下，形成组合的可能性非常低，从而出现的生态位也非常少。但是，随着更多的单个实体的进入，我们就可能会观察到生态位的爆炸性增长。因为可能出现的组合数量的增加是呈爆炸性的，可能产生的生态位数量的增加也是呈爆炸性的。

第二，如果新实体对旧实体的替换发生在等级体系的基底附近，那么系统复杂性坍塌的概率就会非常高，因为这种等级体系的每个层级都依赖更基础的层级。而如果发生在等级体系的端点附近，那么复杂性坍塌的概率就会很低。因此，在一个共生进化系统中，多元化是扩展还是

坍塌，在很大程度上取决于等级体系的依赖结构。

第三，这里存在两种正反馈，或者说循环因果关系。这是这种机制所固有的。新实体的产生可以强化进一步的新实体的产生。原因很简单，因为系统中出现了可用于进一步的“自适应辐射”的新“遗传材料”。新实体的出现还为其他新实体的出现提供了生态位。反过来，这也就意味着，当只有很少的新实体被创建出来时，也很少会有进一步的新实体出现，因而也就几乎不会创造出新的生态位。因此，系统在这种情况下将呈现基本静止的状态。当新实体迅速出现时，新的生态位将快速增加，从而进一步产生新的实体和新的生态位。然后系统就可能经历一种类似“寒武纪大爆炸”的快速增长。因此，我们可以预期，这样的系统可能在很长时间内处于相对静止的休眠状态，但是偶尔也会突然进入复杂性爆发式增加的时期。也就是说，我们预测，它们会经历间断均衡。

这种机制，即系统的复杂性随着新的生态位或利基的创造而增加，是我们大多数从事复杂性研究的人很熟悉的。例如，斯图尔特·考夫曼就写过很多论著，介绍了各种各样的自强化多样性的例子。然而奇怪的是，在传统的生物学文献中，却很难找到有关这种机制的讨论。博纳（Bonner）在1988年出版的《复杂性的进化》（*The Evolution of Complexity*）一书中也没有提到它，尽管他用了整整一章的篇幅来讨论以多样性形式表现出来的复杂性。沃丁顿（Waddington）认为，随着有机体多样性的增加，生态位会变得越来越复杂，这种说法有点接近我们这里讨论的这种机制。他还说，生态位越复杂，来填补生态位的生物就会越复杂，而这反过来又会增加生态位的复杂性。但是，他在谈到这些时，所指的似乎是内部结构复杂性的螺旋式上升，而不是共生进化生态的多样性。不过，有意思的是，早在100多年前，达尔文就提到了这种机制，或者至少是与之有密切关系的某样东西。在他笔记的第422页，达尔文这样写道：

> 世界上，动物的极大数量取决于它们的结构和复杂性……因此当形式变得很复杂时，它们也就打开了增加自身复杂性的手段的大门。[1]

但是，这种说法同样可以理解为与内部结构复杂性有关，而不是与生态多样性有关。

机制2：结构深化

我将导致复杂性随时间增加的第二种机制称为结构深化（structural deepening）。这个机制适用于单一实体，包括系统、生物体、物种及个体，它们都在某个“背景”下进化。在此，我们可以将这个背景视为它们的“环境”。通常来说，在竞争的强大压力下，这些系统会运行在自己的性能极限之内。但是，它们也可以通过增加某些功能或子系统，来打破这些性能限制。增加了这些功能或子系统后，它们就能够：（a）在更广阔的或更极端的范围内运行；（b）感知并对异常情况做出反应；（c）服务于其他系统，使其运行得更好；（d）提高其可靠性。通过这种做法，它们增加了它们的“结构深度”或“设计的复杂性”。当然，在加入这些功能或子系统后，系统的运行仍然可能达到其性能极限。在这种情况下，还可以通过进一步添加子子系统，来再次突破这种限制。循环反复地推进上述过程足够长的时间后，原来的系统就会被更深层次的功能和子功能所“覆盖”，从而可以极大地改善自己的运行性能和适用范围。当然，这样一来，它内部的结构也就变得更加复杂了。

技术进化的历史提供了许多结构深化的例子。例如，由弗兰克·惠特尔（Frank Whittle）和汉斯·冯·奥海因（Hans von Ohain），在20世纪30年代分别独立设计了最早的燃气轮机和喷气式航空发动机，就是一个很好的例子。这种发动机的工作原理是：压缩进气，在压缩空气中点燃燃料，并通过涡轮机释放燃爆的混合物以驱动压缩机，然后以极高的速度排出空气来提供推力。惠特尔的原型有一个可移动的部分，即压缩机-涡轮机组合。但是，多年以来，在来自商界和军界的利益团体的推动下，巨大的竞争压力导致航空发动机不断做出改进。设计师们不得不想方设法，让航空发动机能够克服极端的压力和温度所带来的限制，并能够应对各种突发的特殊情况，有时是通过使用更好的材料，但更多的时候是通过添加子系统。

因此，随着时间的推移，最终的成果是，通过使用不止一个压缩机，而使用一组压缩机，即压缩机的系统，实现了更高的空气压缩比。其他一些改进包括：采用可变位置导向叶片控制系统，以便在高海拔和高速度条件下吸入更多的空气，同时降低发动机熄火的可能性，从而提高了效率；增加排气阀控制系统，以便在气压达到某一临界点水平时，将空气从压缩机中排出，这样就能够减少发动机失速的可能性；添加二次空气流系统以冷却红热的涡轮叶片，并对集液槽腔加压以防止润滑剂泄漏，涡轮叶片也通过一个在内部循环空气的系统来冷却；在军事应用中，为了在空中作战条件下提供额外的推力，增加后燃烧器组件；为了减少发动机起火的可能性，增加复杂的火警探测系统；为了防止进气区

域中积聚冰块，添加了除冰系统；另外还有增加专用燃料系统、润滑系统、可变排气喷嘴系统和发动机启动系统。

但所有这些系统都需要进一步的子系统，以便监视和控制它们，并在它们达到极限时提高它们的性能。而增加的子系统，又需要更进一步的子子系统来增强它们的性能。现代的航空燃气涡轮发动机的功率，比惠特尔的原型高出了30倍到50倍，而且非常复杂。但是，惠特尔发明的原始的简单系统还在，只不过被子系统叠子系统地覆盖在了一个极其复杂的互连模块和部件的阵列当中。现代航空喷气发动机的部件超过了2.2万个。[2]

以航空发动机为例。在这种机制中，在持续的竞争压力的推动下，设计师们必须不断地将新功能和新模块添加到系统中，以突破原有的性能限制和处理异常情况，或适应更加复杂的环境，从而导致复杂性不断增加。行文至此，读者应该已经看得很清楚了，通过结构深化而使复杂性不断增加的这种机制，不仅适用于技术，而且适用于生物有机体、法律制度、税收体系、科学理论甚至连续发布的软件。

约翰·霍兰德的遗传算法为我们提供了一个实时观察结构深化的“实验室”。在“候选解决方案空间”运用遗传算法搜索可行的解决方案的过程中，会先出现一个大致正确的，或者用霍兰德本人的术语来说，“粗糙的模式”。它可能只比它的“竞争对手”稍稍好一点。但是随着搜索的继续进行，更加好的解决方案也开始出现。这些解决方案具有更深的结构，更精细的子模式，是对原始解决方案的“精炼”，能够处理异常情况或克服原始解决方案的某些限制。最终的解决方案或模式的组合，可能在结构上很“深”，也很复杂。在运用遗传算法搜索解决方案的过程中，还可以观察到结构深化的逆转。这种情况发生在相当长的一段时间内，一直占据了主导地位，并被精炼得相当充分的“粗略模式”被另一个新“发现”的、有所改进的“粗略模式”取代的时候。这时，依赖于原先那个粗略模式的子模式的层级结构就会坍塌。于是，对好的解决方案的搜索开始集中于新的模式，它们将会得到精炼。这种情况在遗传算法的搜索过程中可能会发生很多次。

约翰·科扎（John Koza）的遗传编程算法也提供了一个类似的“实验室”。科扎让一些代数表达式在他的“实验室”中进化，最终的目标是让它们解决给定的数学问题。在科扎的实验中，我们可以观察到，越来越深的“结构深度”被构建到了当前性能最佳的代数表达式中，用来描述表达式的算法解析树所长出的诸多分支。

图9-1给出了在科扎的实验中，结构随着对“好的解决方案”的搜索的进行而不断增长的一个例子。正如我们很容易就可以看到的，这种机制也不是单向的。当新出现的表达式替换掉了原始系统中“根基”附近的符号表达式时，就会发生结构深度和复杂性的反转。总的来说，结构深度是增加了，但是“在途中”会间歇性地反转成为相对更简单的结构。

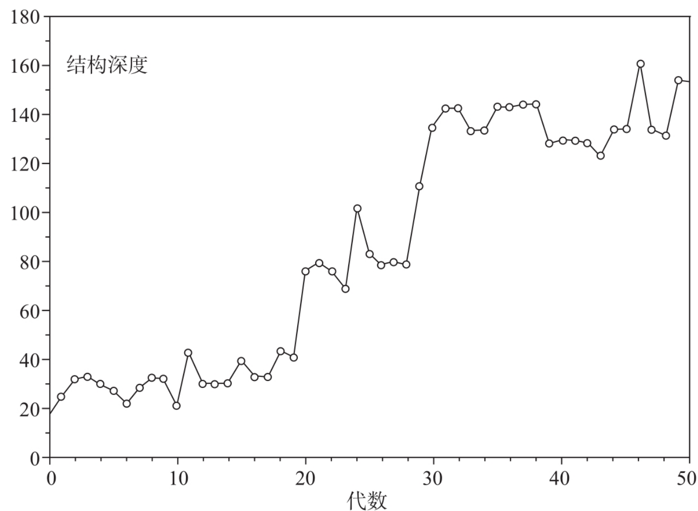


图9-1 当前最佳表达式的结构深度与搜索代数之间的关系

复杂性在一个系统的“根基”附近坍塌这种现象，在一个大相径庭的情境下也可以观察到。科学发展史上，当新的理论突然取代精心构筑且得到了详细阐述的旧理论时，这种情况就出现了。其中一个很好的例子是，经开普勒-牛顿改进过的哥白尼理论造成了托勒密天文学的崩溃。这种新颖的理论体系，只需要寥寥几个简单的定律，就可以解释行星的运行轨道，从而摧毁了巨大而复杂的托勒密体系的根基。由于新的理论的解释力胜过旧理论百倍，托勒密体系也就从此一蹶不振了。类似地，在20世纪30年代，推进原理非常简单的惠特尔式喷气航空发动机相当彻底地取代了活塞式航空发动机。当时，为了克服在高海拔地区非常稀薄的空气对高速内燃发动机的限制，后者已经不可避免地变得过于复杂了。在很多其他不断进化的系统中也是这样。突然涌现的简单性，往往会切断原先的复杂性增长之路，并建立一个全新的基础，让复杂性再次增长。总的来说，在这种复杂性与简单性之间的“拉锯战”中，在更多时间内占据上风的通常是复杂性。

到目前为止，我已经阐述了两种明显不同的机制。第一种机制是，生态系统，即许多个体的集合，在进化过程中变得越来越复杂、越来越

多样化。第二种机制是，生态系统内的个体在进化过程中，内部变得越来越复杂、结构变得越来越深化。当然，在许多系统中，这两种机制是同时发挥作用的，而且它们会相互作用、互相交替，甚至彼此竞争。

这一点在克里斯蒂安·林格伦对博弈论的策略进化的研究中，可以很清楚地看出来。林格伦构建了一个基于计算机的模型，让各种策略为获得收益而相互竞争。这是一个有限的多次重复的囚徙困境博弈，策略之间的配对是随机的，博弈是一对一进行的。这些相互竞争的策略用编好码的位串来表示。在位串中的“位”代表对先前行动的记忆，而行动是由策略决定的。策略偶尔会发生变异。在这种共生的环境中，成功的策略将迅速繁衍，而不成功的策略则会消失。在林格伦的世界中，我们可以清楚地观察到，随着新的、不断进化的策略创造了更新的策略可以利用的生态位，策略的多样性会不断增加。这种情况与我在上面讨论过的第一种机制相当类似。但是与此同时，策略本身也变得越来越“深”，在这里“深度”指代码串或记忆的长度，竞争中得到的回报也随策略变得越来越微妙而上升，这就是符合第二种机制的情况。事实上，这两种机制是相互作用的，一个新的、成功的、更“深”的策略一旦出现，就会使许多以前的、更加简单的策略消失。这样一来，多样性坍塌了，随多样性而来的许多利基也不复存在。但是在这之后出现的又是这样一个阶段：更新的、更“深”的策略突变出来并得到增殖，从而使得多样性再次增长。因此，新的、更“深”的策略就可以破坏旧的多样性，也可以在新的、更“深”的策略当中引发新一轮的多样性增长。这样一来，共生多样性的增长以一种偶发的方式，随着策略的结构深度的增长而交替变化。在生物进化史上，不难找到明显类似于这种过程的例子。例如，许多生物学家认为，多细胞、真核生物的出现意味着生物系统的“深度”的增加，而且也促进了距今6亿年以前寒武纪的多样性爆发。

机制3：捕获软件

接下来讨论复杂性的增长的第三种机制。这种机制与前两种机制完全不同。事实上，这种机制基本上只与复杂性的快速涌现有关，而与缓慢增长的复杂性没有太大的关系。它刻画的是一种我称之为“捕获软件”的现象。这是“外向型系统”（outside system）出于自身目的，通常是信息方面的目的而“接管”系统外部的比较简单的元素，并让它们“完成任务”的行为。在这种机制中，通常而言，“外向型”系统会“发现”外部的一些较简单的元素，并“认为”可以将它们用于某些基本目的。这些元素

则通常“具有”一组规则，它们决定了它们如何被组合和被使用，这种规则可以称为“交互式语法”。这种语法通常允许简单元素以许多种方式组合起来，而且当“外向型系统”学会这种语法后，它还能学会利用组合中的这些元素来实现自己的目的。在这种学习非常成功时，“外向型系统”能够学会运用这种交互式语法，来对这些简单的元素进行“编程”，并且在复杂组合中利用它们来实现自己的多项目标。

初听起来，这种机制似乎有点奇怪，而且许多人都会觉得有些陌生。因此，我在这里先举一些例子来说明一下。一个非常简单的例子是电子技术。在过去的几个世纪里，我们人类已经学会给电子“分配任务”，让电子完成诸如传输声音和视觉信号、控制复杂机械和进行计算等“工作”。最初，在法拉第和富兰克林的时代，人们对电子和静电的工作原理的了解非常有限，所以电子的“用法”也很少。但是在19世纪和20世纪的前几十年里，我们开始学会了电的“语法”，即一组包括导电性、电容和电阻的操作规则，它们可以控制电子的运动、放大电流。自那之后，我们也就慢慢地学会了“捕获”电子，并对电子进行“编程”，以服务于我们自身的目的。在这个例子中，上面提到的“简单元素”指的是电子，“外向型系统”指的是我们自己，即人类用户，“语法”指的是电磁学定律，“可编程输出”指的是电子设备所用的各种技术用途。在输出的层面上，这里存在着迅速的“适应”。我们将电子作为“可编程软件”来使用时，各种技术的用途都会迅速改变和迅速扩大。但是在语法和载体层面上，在这个例子中是不存在“适应”的。这是因为，电和电子的“行为”是由物理定律“规定”的，而物理定律至少在人类时间的范围内是不会改变的。

不过，在其他一些使用“捕获软件”的场合，交互式语法却不是不可改变的，而是可以在“捕获”软件的过程中自己更新和进化的。这方面的一个例子是人类语言的进化方式。原始人类可能在距今几十万年前，就学会了发出一些粗糙的声音，用于满足提出警告、表达情感，以及其他简单的需求。然后，非常缓慢地，在比较近的进化时间尺度上，语言的基本规则，即语法开始涌现出来了，从而人类可以将声音组织成简单而连贯的表达。最终，经过数千年的进化，这些声音或音素加上语法，终于进化成了一个复杂的交互系统：一种语言。语言可以被“编程”，用于构成能够表达含义非常细致微妙的陈述、问询和命令。

在这个例子中，发挥着载体作用的“简单元素”，是人类说的话中的声音或音素。“外向型系统”是人类社会，它“捕获”这些简单元素，使它

们成为一种软件、一种语言。“语法”是一个句法体系，它确保意义的一致性和共通性。当然，人类并不存在一种共同的语言，因此也没有单一的句法结构。语法必定是通过社会惯例的缓慢进化而涌现出来的，而对语言效率和一致性的需求，以及人类大脑中组织语言活动的方式，都会使语法受到一定限制。当然，作为对语言进化的一种回应，人类的发声器官和大脑的生活解剖结构也会发生变化。总而言之，从这样一个进化过程中涌现出来的语言是一种“可编程软件”，我们可以把它的潜在输出视为语言可以表达的所有有意义的句子或语句的集合。

在语言这个例子中，适应可以发生在所有层面上。在程序输出层面上，适应是瞬时完成的。我们可以把人们说出的一个句子视为一种一次性的、极其快速的软件输出层面上的适应，其目的是适应人们用这个句子想沟通的东西。而在语法层面上，适应意味着语言本身的变化。这通常会以漂移的形式出现，其发生是一个缓慢的和连续的过程。这是因为，语法的任何突然改变或大的偏差，都会使当前的“输出程序”无效。而在音素或简单元素层面上，适应或者说改变和漂移是最慢的。在这种载体层面上的任何轻微变化，如果不是极其细微和连续的，都很可能导致整个系统的表达的失败。非常缓慢的漂移确实可能发生，如元音在一代人或两代人的时间内可能会有所改变。但是总的来说，存在着一种强有力的机制，它可以保证载体元素的行为方式保持恒定。

捕获软件这种机制的一个特别有意思、特别有说服力的例子是，近年来在金融市场中复杂的衍生产品的出现和广泛使用。在这个例子中，“外向型系统”是金融界。最初，金融市场上交易的是有价值的东西：大豆、证券、外汇、市政债券、第三世界债务、抵押贷款、欧元，以及任何可以持有的有价值的东西。这些东西的价值可能会出现波动。用金融术语来说，它们是“基础资产”，或者被称为“基本面因素”。它们构成了我现在所要考虑的系统的简单载体元素。

在这样的金融市场的早期，基础资产通常是被简单地持有，或以其内在价值被交易的。但是，随着时间的推移，一种新语法形成了。一些交易者发现，他们可以：（a）创设一些与或有事件相关的期权或选择权，这种或有事件会影响基础资产的价值；（b）就像股票指数一样，将几种基础资产合并在一起来创建一个相关指数；（c）发行期货合约，约定在某一时间，如未来的60天或一年，交付或取得基础资产；（d）发行以基础资产为担保的证券。但是请读者注意，这些“衍生工具”与或有事件挂钩的期权、指数、期货和证券，本身也是价值的元素，

因此它们也可以成为“基础资产”，拥有自己的交易价值。金融市场可以再次对这些新的“基础资产”，应用上面的（a）、（b）、（c）或（d）。然后，我们就有了证券期权、指数期货、期货期权以及证券指数期权等进一步的衍生产品，它们都可以进行交易和掉期。

有了这样一种语法，衍生产品专家就可以对这些元素“编程”，然后组成一个“包”了。这种“包”是为那些有非常特殊、非常复杂的财务需求的客户量身定制的，为他们提供了他们想要的融资、现金流和风险敞口的组合。当然，金融市场不是在一夜之间发明这样的“编程方法”的，它是在若干个市场中半独立地发展起来的。一开始只利用了简单的载体元素，然后再与自然的金融语法相结合。

从我在上面给出的这些例子来看，使用和捕获简单元素、用于实现自身目的的系统似乎总是我们人类社会。但是，事实当然并非如此。下面，我再举两个来自生物领域的例子。第一个例子是神经系统的形成。随着生物的进化，某些生物体开始给一些专门的细胞“分配任务”，即让它们只服务于感觉和调节对外部刺激的反应这个简单目标。这些专门的细胞又反过来发展出了自己的交互式语法。然后，整个有机体又使用这种语法，来对这个互连的神经系统进行“编程”，以便利用它来实现自己的目的。类似的一个例子是免疫系统的形成。人们发现，高等生物的免疫系统中的细胞的祖先，最初只用于一些简单的目的。随着时间的推移，这些祖先细胞也开发出了有用的互动规则，即交互式语法，并最终发展成为一个高度可编程的系统，它可以保护有机体免受外部抗原的侵袭。

对于生物生命本身，也可以用这种方式来思考。不过，这里的情况比前面的那些例子要复杂得多。生物有机体是由一系列主要是细胞构成的模块所构成，而细胞又是由相对较小且种类有限的、大约50种的分子构成的。这些分子是相当简单的，在所有陆地生命中都普遍存在，它们是生物构建的载体。而且，这些分子只须运用一种相对简单的语法，就可以组合成适当的结构。这种语法包括相对少量的代谢化学通路，是一种代谢语法，它又受一些酶的调节，或者说被一些酶编程。当然，进行这种“编程”的这些酶并不是有意识的、有目的性的，它们本身又是第二层的编程系统中的载体。它们要受一种复杂的基因表达“语法”调制。基因表达语法会根据它们所在的生物体状态的反馈，打开或抑制编码它们的基因或DNA中的繁殖“开关”。于是我们就看到：前一个捕获软件系统，是受后一个捕获软件系统调节的。前者通过蛋白质或酶对简单代谢

途径进行编程，以形成和维持生物结构，后者通过核酸根据生物体的当前状态，对蛋白质或酶进行编程。

在上面这个例子中，整个系统是封闭的，是不属于对某种生物对象进行编程，以实现自身目标的那种“外向型系统”。在短期内，每个有机体根据自身当前的发育状况和当前的需求，来对自身进行编程。而从长期的角度来看，整个系统，即那些生存下来并相互作用和共生进化的生物体的生物圈以及各种环境和气候因素会成为编程者，并以基因序列集合的形式“写好”代码。当然，基因序列任何时候都将生存和存在。只要没有一个“外向型系统”，我们就不能说这些可编程的系统曾经被“捕获”。相反，它们是涌现出来的，而且会自我引导，或以自展方式开发出载体、语法和软件。从这个角度出发，生命的起源可以解释为一个由物理系统承载的软件系统的涌现，即一个能够学会对自身编程的可编程系统的涌现。

在上面所有例子中，捕获软件都是一个非常成功的进化策略。在这种机制下，系统只须通过对捕获的系统重新编程，来形成另一种不同的输出，就可以非常快速地实现适应。但是，因为语法和载体的变化会破坏现有的“程序”，因此我们通常会希望它们是被“锁定”的，或者即使有变化，也是越缓慢越好。这就解释了为什么遗传序列可以轻易地改变，但是遗传密码却不会；还可以解释为什么新生物体可以出现，但是细胞和代谢化学通路则是相对固定的；还可以解释为什么新的金融衍生产品不断涌现出来，但是证券交易所的规则却会保持相对恒定。[3]

复杂性的坍塌

在本章中，我阐述了复杂性随进化而增长的三种机制。这三种机制是：通过自我增强的多样性的增加，或者通过结构复杂性的增加来突破性能限制，或者通过系统“捕获”更简单的元素，并学会将它们“编程”为“软件”来用于它们自身的目的。当然，我们不能指望这种复杂性的增长是稳定的。相反，在所有这三种机制中，我们都预测复杂性的增长是间断性的、世代性的。我们也不能指望复杂性的增长是单向性的。前两种机制肯定是可逆转的，所以我们会预期，复杂性的坍塌随时都可能随机地发生。

在我们更加深入细致地对进化进行研究时，我们在生物领域之外发现了大量我们感兴趣的例子。任何一个系统，只要它的结构要面对提升

自身性能的压力，并有一个“继承下来的、可变的结构的世系”，都会呈现出进化现象。我相信，能够证明复杂性与进化之间的联系的例子会不断地涌现出来。它们不仅来自生物学领域，还来自经济学、自适应计算科学、人工生命、博弈论等领域，在这些领域中能找到复杂性与进化之间的联系。重要的是，本章描述的这些机制，适用于所有这些进化情境中的任何一个例子。

COMPLEXITY AND THE ECONOMY

10

认知科学

打开经济学黑箱的金钥匙

当人类行为主体面对复杂的或不确定性的决策问题时，他们在推理时所运用的不是演绎理性。那么，他们用的是什么呢？认知科学告诉我们，在这种情况下，我们会“联想地”进行思考：我们从经验中找到类似的情境，并用这些情境去拟合我们所面对的问题，然后从中得到一些启示。

大多数经济学家都认为，当人类行为主体面对复杂的决策问题，或者面对包括了具有根本不确定性的决策问题时，他们在推理时所运用的不是演绎理性。那么，他们用的是什么呢？认知科学告诉我们，在这种情况下，我们会“联想地”进行思考：我们会从自身的经验中找到类似的情境，并用这些情境去拟合所面对的问题，然后从中得到一些启示。现在这篇文章探讨了这种类型的推理对经济学的意义，并提出了在经济学

中对这种推理建模的若干方法。这篇文章还强调，各种情境的记忆和经验，对于我们的推理是必不可少的。有鉴于此，学习经济学的学生应该深入地钻研经济史，而不能仅仅关注经济学理论。

这篇文章最初是在2000年出版、由戴维·科兰德主编的《复杂性视野与经济学教育》（*The Complexity Vision and the Teaching of Economics*）一书中发表的。

* * *

伯特兰·罗素（Bertrand Russell）在自传中告诉我们，他在学习了半年的经济学后，就对这门学科失去了兴趣，因为他认为它实在太简单了。马克斯·普朗克也在中途放弃了经济学，可是他的理由却是因为它实在太困难了。至于我自己，之所以会成为一名经济学家，是因为我接受过数学方面的专业训练，而且我一开始也像罗素一样，觉得经济学很容易。几年之后，我就从罗素的立场转换到了普朗克的立场。从其本质上来说，经济学是非常“难”的。在本章中，我将详细解释我为什么会得出这样的一个结论。

经济学从本质上说到底是难是易，取决于提出经济问题的方式。如果在构建一个经济问题时假设决策是理性的，那么通常而言都可以求得一个确定的“解”。这种经济学是简单的，无非是从问题跳到问题的解而已。但是，在这种经济学中，行为主体到底如何从问题得到问题的解的，却仍然是一个黑箱。而且，行为主体是否真的能够到达那个解，也是无法保证的，除非我们能够打开这个黑箱看一看。但是一旦我们试图打开这个黑箱，经济学马上就会变得非常困难。

曾几何时，经济学家们认为，我们对问题与问题的解之间的联系的假设是有道理的。在一篇现在已经很著名的文章中，拉斯特讲述了威斯康星州麦迪逊大都会公共汽车公司维修主管哈罗德·泽克（Harold Zurcher）的故事。在整整20年里，泽克一直负责为他的公司制定更换公共汽车发动机的计划。这个复杂的问题需要他平衡两个相互冲突的目标：最小化维护成本与最小化发动机意外故障率。拉斯特用随机动态规划方法找到了这个组合优化问题的解，并将这个解与泽克的优化方案进行了对比，结果发现两者之间拟合得相当好。据此，拉斯特得出了这样一个结论：尽管面对的是一个非常复杂的问题，但是哈罗德·泽克还是找到了问题的解，因此经济学家的假设，即个体能够找到复杂问题的最优解并不是一个坏的假设。

泽克的例子给我们留下了一个很大的问题：个体能够找到经济问题的最优解这个假设，是不是合理的。如果是合理的，我们是否就可以不用去研究决策过程的细节了？在简单的情况下，答案是肯定的。然而，在大多数情况下，答案却是否定的。读者不妨想象这样一个“决策问题之海”，它包含了所有我们感兴趣的、有明确定义的经济问题。这些问题构成了海水浅表层，复杂的决策问题构成了海水的深层和底层，海水越深，问题就越困难。海平面附近是类似于“井”字游戏的那些问题，下面是跳棋一级的问题，更深的是国际象棋和围棋一级的问题……在理论上，我们可以说，国际象棋的“解”是存在的，它将采取纳什混合策略的形式，但是我们不能保证人类行为主体肯定能得到这种“解”。因此，可以求得“解”的、像“井”字游戏这样的问题，只存在于海平面及海平面以下半米左右的海水内。而在比这更深的海水处的问题，则无法保证有解。我们还可以进一步列出这些行为主体要面对的许多更加复杂的问题。但是，大多数问题是没有明确界定的。泽克面对的问题恰好位于经济主体可以通过“理性”求解的问题与不能通过“理性”求解的问题之间的分界线上。若“海水”比它更深一点，那么经济问题的“解”就无法与“理性”匹配了，甚至可能根本不存在。

那么，在这些更深的层面上会发生什么呢？在现实生活中，人类决策者绝不会因为一个问题很难或尚未明确界定就望而却步。也许，当问题太复杂以致无法求解时，或者当问题尚未得到明确界定时，我们应该说行为主体面对的不是传统的问题，而是一种情境或情况。他们必须处理好这种情况，他们必须自己框定（framing）问题。从许多方面来看，这种框定正是决策过程的最重要的部分。在考虑如何框定问题时，你必须考虑问题和要采取的行动之间的关系。介于问题与行动之间的是认知，而且在问题与问题的解之间有很多很多东西，只要将这些东西纳入考虑的范围，那么经济学就会变得非常困难。关于介于问题与行动之间的认知问题，我可以将它转换为如下几个问题：人们如何理解一个问题的意义？个体如何处理更加复杂的问题？我们怎样才能真正认识一些东西？

在本章中，我想从认知心理学家的角度来考虑认知问题，并利用得到的结果来讨论以下两个问题：经济建模和经济学研究生教育。

心智是什么

在经济学中，我们一直秉持着一个简单但陈旧的心智概念。心智（Mind）被视为一个保存数据的容器。数据通过与世界的互动不断更新，而心智则根据这些数据进行演绎推理。所有这一切在经济学中都是隐含的假设，因为在经济学中我们不会谈论“心智”。但是，我们确实认为心智或“那个推理机”，是在数据集的基础上进行演绎推理的。在经济学理论中，这种观念反映的是，将对世界的信念视为以当前数据，即当前信息为条件的变量的期望值，以及基于这些信息来制定解决方案。这是一种“速记”，也是一种合理的抽象，各门学科在许多情况下都可以利用这种方法，而且效果不错。但是我们必须超越它，只要我们来到了比“决策问题之海”海平面低半米左右的地方。

因此，我们应该从更加深层的观点，即认知科学的观点，来讨论心理和认知过程。请读者想象一下，在某个静谧的晚上，你正在读一本小说，或者就拿哈尔多尔·拉克斯内斯（Halldór Laxness）所写的《独立的人民》（*Independent People*）来说吧，你很喜欢这本小说。表面上看，这似乎是一件再平常不过的事情了，但是实际上却很复杂。印在纸页上的那些黑色和白色的标记，会聚焦到你的视网膜背面的光传感器或“像素”上。这些感官知觉随后会传递到你的大脑的后部，并映射到特定的视觉神经结构。然后，不知怎的，字母和单词被解析出来。同样不知怎的，这些字母和单词通过对句法的理解而结合在了一起。我在这里说“不知怎的”的意思是，认知科学家不知道在这个知觉过程中，发生这些事情的确切机制是什么。再然后，意义不知怎的从句法结构中涌现出来了。但是，“意义”到底是什么意思？在这个例子中，意义指的是一组关联或联想。你可能刚好读到了一句描写雨景的话：“雨一直下，平稳地、执着地，把整个郡都笼罩起来了。它落在衰朽的沼泽草上，落在那个‘是非不断’的湖上，落在铁灰色的砾石砌成的地上，落在了俯视众多小山的阴沉沉的大山上……把所有的风景都玷污了。”这些词会触发联想，事实上是触发相关的记忆。然后，你的脑中会形成一副画面或一组画面。这些相关的记忆和画面，反过来又会触发你所称的“情感”或感觉。感觉往往是很微妙的，在拉克斯内斯的世界中当然也不例外：雨的阴暗、砾石铺就的路的泥泞、山的压迫感、小农场的潮湿气味，这些感觉确实非常微妙，它们其实是我们智慧的一部分，是我们认知的一部分。它们是我们赋予符号意义的一部分。阅读并理解读到的内容的含义，涉及相关的记忆和相关的感觉。这种理解何以成为可能？认知科学家们迄今仍然未能很好地解释这一切是如何发生的，因此法国思想家亨利-让·马丁称之为“神秘的炼金术”。

对于这种炼金术，普林斯顿大学认知心理学家朱利安·杰恩斯（Julian Jaynes）这样描述道：

> 啊，这是怎样的一个世界啊！看不见的多姿多彩，听得到的无声无息。这个无形无质的心智国度竟是如此神奇！啊！那不可窥见的本质！那无法触及的记忆！还有那难以预测的遐想！它莫非也要保护它的隐私吗！这是一个秘密剧院，上演的戏剧中充满了无言的独白和抢眼的秘密。这是一幢看不见的豪宅，收藏着所有的心情、冥思和奥秘。这是无限的失望和无穷的发现的胜境。它是一个王国，在这里，我们每个人独立地推行自己的统治，质疑我们想要做的，指挥我们做自己能做的。它是一个隐秘的道场，我们在这里一边研习秘籍，一边反思自己做过的、谋划自己未来要做的。它是一台“内部对讲机”，但所说的，更多的是关于“我自己”，而不是关于任何我可以在镜子中找到的东西。意识就是自我本身，它无所不包，但又什么都不是。那么，它到底是什么？它来自哪里？它的意义何在？

我在这里想强调的一点是，从书中抽象出来的意义并不在书本中，而在我们的心智中。对于这一点，哲学家们，如康德在18世纪时就已经开始认识到了，但是直到20世纪还没有完全阐明。我们通过我们所建立的关联来构建意义。

如果读者觉得这个论断听起来有些奇怪，那么请想象一下：让一个俄罗斯人看陀思妥耶夫斯基的俄文版小说《独立的人民》，与让一个不懂俄语的外国人看同样的一页纸，会有什么不同。他们两个人获得了完全相同的数据，但是那个俄罗斯人能够在他看到的那些西里尔字母与意义之间建立起关联，从而使他自己获得关于那些书面文字的感觉数据“活”起来。而那个不懂俄语的外国人虽然看到了完全相同的数据，却无法建立起关联，因而那些字母对他没有意义。

从这个角度来看，意义是“强加”的，它因我们强加的关联而得以涌现出来。是我，而不是陀思妥耶夫斯基这个人，也不是《独立的人民》这本书给我带来了意义。认为后两者给我带来了意义，那只是一个错觉。是我给《独立的人民》带来了意义。是我在理解，在赋予它意义。是我给我看到的东西强加了各种关联，给我看到的东西强加了意义。而且，不仅是某种陈旧的意义，还有刚刚从这本书与我的神经记忆的关联当中涌现出来的意义。

下面，我再给读者举一个例子，因为我想从这个要点锤炼出更多东西。爱尔兰诗人叶芝有一首著名的小诗这样写道：“我的爱人和我确曾相会在柳园下边，她那一双雪白的小脚款款走过柳园……她让我从容看待人生，如堰上长青草，可我，那时年少无知，如今悔泪滔滔。”[1]这些诗句对不同的读者会产生不同的影响，即对不同的读者有不同的意义。

读者不妨问一下自己，当你从河边走过，看到河中的堰时，它对你有什么意义。对于我来说，它有“重大”的意义，因为孩提时，我和我的朋友们经常在各种堰的附近玩乐嬉闹。堰是河流中的小坝，通常覆盖着碧绿的尾草和各种藻类。当然，我也知道什么是“柳园”（salley garden）。但是，那些不熟悉爱尔兰的人的感受可能会完全不同。他们可能会觉得奇怪，这里说的“柳园”到底是什么啊。也许一个名叫莎莉（Salley）的人有一个花园。也许有什么东西叫“Salley Garden”，莫非诗中的男女主人公在都柏林附近有一栋别墅叫这个名字？在不知道“Salley Garden”到底是什么东西的情况下，你在脑海中可能会浮现出一个美丽花园的形象：繁花似锦、绿草如茵，还有许多园丁在辛勤地劳作着。但是，不是这样的！在盖尔语中，“Salley”这个词要说成“s-a-i-l-e-a-c-h”，意思是“柳树”。因此，叶芝说的那个场景中有柳树，因此旁边不远处应该有水。如果还有一个堰，那么就应该是一条小溪或小河。一旦建立起了这些关联，一切就立刻全都改变了。总之，我的观点是，不同的意义可以强加在相同的数据之上。源于不同关联的意义可能完全不同。

数据，无论是文学中的数据，还是经济学中的数据，都没有固有的意义。它们能够获得的意义，只能是我们给它们带去的意义。不同的人，因为各自拥有不同的经验，肯定会建构出不同的意义。

心智是快速的模式完成器

联想引擎（**Associative Engine**） 哲学家和认知科学家安迪·克拉克提出的一个概念。现代认知心理学告诉我们，我们的大脑就是一个“联想引擎”。通过联想，我们给事物强加

了可理解的模式。

那么，现代认知心理学从这些例子中得出了什么结论呢？第一个结论是，我们的大脑是一个“联想引擎”。这个术语是由圣路易斯华盛顿大学的哲学家和认知科学家安迪·克拉克首创的。我们非常擅长联想，就认知而言，我们可以说联想几乎相当于我们所做的一切。通过联想建立关联，我们给事物强加了可理解的模式。这里不妨借用克拉克首创的另一个术语：我们是快速的模式完成器（pattern completer）。

举一个例子，如果我看到墙角露出了一条尾巴，它是黑色的、毛茸茸的，还摇个不停，我就会说：“那里有一只猫！”但是，实际上是一个小男孩，在一根棍子上装了一条尾巴，试图欺骗我。我刚看到的时候却不会这样去想。我的心智天生就注定了，它不会让我这样去想的。如果我强烈怀疑那应该不是一只猫，或者如果我看到了一个搞恶作剧的小男孩，那么我可能会说：“哦，那是一只猫或一个小男孩。”但是在没有看到任何一个小男孩的时候，我肯定只会说：“嘿！我看到了一只猫。”尽管我看到的并不是一只猫，而只是看到了一条黑色的尾巴。著名哲学家伯特兰·罗素讲过的一个故事也反映了同样的观点。一个小学生、一个牧师和一个数学家从英国乘坐火车到苏格兰去。小学生从车窗往外看风景时，发现了一只黑羊，于是他就说：“哦！看哪！苏格兰的羊是黑色的！”牧师听到这句话后，就说：“你说的不对。严格地说，我们只可以说苏格兰有一只羊是黑色的。”数学家却说：“这样说还是不准确。我们可以说的是，我们知道，在苏格兰至少有一只羊，至少有一侧是黑色的。”

认知科学反复“告诫”我们，我们人类在日常生活的决策中，不会像数学家那样去演绎，而只会像小学生那样去联想。而且更加重要的是，我们这样做有一个非常强大的理由：这种倾向是进化造就的。我们人类祖先在10万年前就拥有了这样的能力：只要嗅一嗅空气，瞬间就能把空气中的湿气与远在几公里之外的水源之间建立起关联。这种能力对生存有非常重大的价值。这里存在着一种“快速完成模式”，即在一瞬间就从最微弱的线索中推测出水的存在的能力和程序，它有助于我们的生存。但演绎逻辑却不是这样的，除了一些最“有空”、对生存最不重要的情况之外，我们根本不会使用演绎推理。

认知心理学家还告诉我们，演绎推理本身也主要是联想性的。例如，在解决球面三角学中的一个问题之前，我要先解决这样一个问

题：“这个问题是球面三角学问题吗？”然后把这个问题与适当的框架建立起关联。在那个框架内，我再把结构和符号与该问题的感觉数据关联起来。我只有通过这样的联想，才能一步步地解决问题，即把各种关联拼合成一个模式。

我这样说，并不意味着人类大脑所能做的一切都是联想，但是从认知的角度来看，联想确实是我们做的主要事情，而且我们在做这件事情时，速度非常快。我们的神经系统会对许许多多的关联进行快速搜索，然后确定某一个为“意义”。偶尔，这个过程也会变慢，因此我们可以观察它是如何“行动”的。例如，几年前曾经相当流行的三维光学幻视图片游戏：让你盯着看一些二维图片，在盯视半分钟之后，一个3D图片就会“蹦出来”。因此，我们的大脑要对大量的联想进行处理，将它们集结成模式，形成各种各样的隐喻。

隐喻就是更加复杂的、带有含义的联想。利用隐喻，我们将“这个”与“这个”、“那个”与“那个”相类比，如果类比是成立的，那么我们就预期会“这样”、会“那样”。隐喻是模式联想的一种形式，我们通过隐喻处理大量信息。总而言之，我们有许多不同形式的联想：形象、记忆、隐喻以及理论。理论，事实上就是得到了阐明的隐喻。在完全起作用时，所有这些东西构成的集合，再加上将它们组合起来的规则（这也是联想）就是我们所称的“心智”。

因此，通过隐喻、记忆、结构、模式和理论，我们的心智特别擅长建立事物之间的关联。换句话说，心智从来不是给定的，它不是一个空空如也的、被动地装载数据的“桶”。心智本身就是涌现出来的。这种观念在西方思维传统中是全新的，但是在东方思维传统中却早就有了很多先例。中国著名理学家程颐和程颢生活在距今大约900年前的宋代，他们都认为心智是“涌现的”。他们没有把心智看作一个容器，而是把心智视为一系列互依互倚、层层相叠的想法。是的，心智不容纳我们的想法。相反，正是这些想法、这些联想包含着心智或构成了心智。心智也决不是固定的，它由联想和控制这些联想的工具组成。正是在这个意义上，心智是涌现的。因此严格来说，我不应该再像以前那样说，意义存在于心智中。因为在认知哲学中，当我们探究得足够深入时，心智概念本身也将消失。意义存在于联想中，我们的神经网络会在呈现给它的数据与自己的联想之间建立关联。

现代认知科学的最新结果告诉我们，我们的立场应该远离如下经济学的传统假设：推理是演绎性的，发生在心智这个容纳各种变量的容器

中，而且这种容器中的各个变量的值是通过“信息”来更新的。如果推理在很大程度上是联想性的，那么就取决于推理者的以往经验。因此，对一个情境的框定，为它构建的意义，都取决于推理者的历史。当然，其结果也必定同样取决于历史。

关于认知，我在这里还要再说一点。有时我们可以粗略地说，只有一个“正确”的意义，即只有一个单一的、正确的联想。但是更加常见的是，即在存在复杂性的情况下，存在多种解释。我们可以采取其中的某一个，也可以采取其中的某几个。通常，如果我们试图解决一个问题，或者试图做出一个决定，如在下棋或打牌的时候，我们会提出许多个假想性的关联，然后对它们进行比较。到最后，当有了足够证据的时候，我们会保留一个或几个。在上面举的“黑色尾巴”的例子中，如果我在几分钟前刚刚看到过一个搞恶作剧的小男孩，那么我就可能会把“猫”和“恶作剧”这两种可能性保留下来，然后等待进一步的证据的出现。

认知过程建模

关于认知，上面已经说了很多了。但是，作为经济学家，必须考虑的问题是，我们怎样才能利用这些东西呢？或者说，在面对复杂的或界定不明确的问题时，我们如何对思维过程建模？

根据如上所述的观察结果，我给出的建议如下。在面对复杂的问题时，作为决策者，经济行为主体必须先找到框定面临的情境的方法。他们会尝试在临时的内部模型、模式或假说与面对的情境之间建立关联来框定问题，并利用这些临时的内部模型、模式或假说。如果他们希望找到行动的指导，他们可能会选择上述模式或模型中的某一个，并进行简单的推理。这种推理仅仅发生在“井”字棋这类游戏的复杂性水平上。当得到了更多的证据后，他们可能会强化或削弱对自己当前的模型或假说的信念。他们也可能在他们的模型或假说不再起作用时舍弃它们，或者在需要时用新的模型或假说替代它们。换句话说，当行为主体面对复杂的问题或无法明确界定的问题时，他们会利用环境提供的线索来形成假想性的模式、框架和关联。这些假想性的模式填补了行为主体理解中的空白。

有了这样一个程序，我们人类就可以处理复杂性问题了。这就是说，我们先构建一些较简单的模型，它们不但是可行的，而且是我们可以处理的。有了这样一个程序之后，我们就能够处理界定不明确的问题

了：在我们面临定义不充分的情况下，我们可以利用工作模型去填补空白。这种行为是归纳性的。它可能会显得有些凌乱，或者看上去有点像“特殊理论”，但是它决不是“理性”或科学的对立面。恰恰相反，它就是科学体系本身运行和进步的方式。

因而在实际的研究中，当面临一个随着时间的推移逐步展开的经济问题时，我们可以建立一个行为主体的集合，他们一般是异质的，并假设他们会联想，这就是说，假设他们能够以心理模型、假说或主观信念的形式建立关联。至于这些信念本身，既可能采取简单的数学表达式的形式，即可以用来描述或预测一些变量或动作，也可能是统计假说，还可能是某种“条件-预测”规则，即如果观察到情况为Q，那么预测结果或行为就为D。这些信念通常是主观的，即不同的行为主体有不同的信念。行为主体在同一时间既可以只考虑一个信念，也可以同时考虑好几个信念，并记住每个信念的“表现”。这样，当需要做出选择时，行为主体根据当前自己最可信的，或者也可能是最有利可图的那一个信念采取行为。他把他的其他信念暂时“藏在自己心底”。说到这里，有的经济学家可能会提出反对意见：行为主体这样做，不就是理性地将自己的若干个假说结合起来吗？得出这样一个结论确实是一个很大的收获。但是，认知心理学告诉我们，我们并不是这样行事的。我们会同时考虑许多假说，但是只会根据目前最合理的那个假说采取行动。而且，一旦采取了动作，那么整体图景就会更新，同时行为主体也要更新对他们的每个假说的置信度。

当然，我在上面提出的这个认知“图式”，本身也是简化和抽象的。但是，它确实刻画了一个重要思想的本质要素：行为主体要将意义“强加”给问题及其情境，或者说行为主体要建立多个框架或信念结构、假说与问题及其情境之间的关联，并让它们相互竞争以此来理解问题。这也是一个存在学习的系统：行为主体通过“学习”获悉他们的假说是否有效。他们的“学习”行为也体现在抛弃表现不佳的假说，并生成新的“想法”而将它们替换掉。请注意，这里存在着一个天生的滞后：行为主体根据他们目前最可信的假说或信念模型来行动，并且在它不再有效时抛弃它，转而采用更好的另一个。某个假说或者关联、信念模型之所以被采用，不是在于它是“正确的”，因为行为主体没有办法知道这一点，而是在于它在过去是有效的。而且，在“值得”将它舍弃之前，必须先积累起它失败的记录。

但是，这里仍然有一个关键问题悬而未决。上面说的那些假说或心

理模型来自哪里？它们是怎样生成的？这涉及心理学中的一个深层次问题，与对象表征和模式识别有关。我不能在这里详细讨论这些。但是就建模而言，还是有一些简单的和实用的选项。有时，我们可以让我们的行为主体拥有“焦点模型”（focal model）。焦点模型是指明显的、简单的和容易处理的模式或假说。这就是说，我们可以先生成一个“焦点模型库”，并分配给不同的行为主体。而在另外一些时候，我们可以给定一个适当的模型空间，然后让某种类型的智能搜索设备，如遗传算法去生成合适的模型。在这里需要提请读者注意的是，无论选择什么方法，我在上面描述的框架都是独立于所使用的具体假说或信念的，就像经济学中消费者理论框架独立于特定的产品选择一样。

这个研究方案真的可以在经济学中付诸实施吗？答案无疑是肯定的。现在已经出现了越来越多的成功例子，如萨金特的研究、“爱尔法鲁”问题、圣塔菲人工股票市场研究。这种类型的研究通常都会发现，必须通过计算来生成“解”、信念及基于这些信念的行动的模式，因为异质信念会使复杂性急剧增加。这些研究通常还会发现一个更加丰富的世界，一个“心理世界”。在这个世界中，关于被研究的问题的各种信念形成了一个生态。有时，这个生态会收敛于某个标准的信念均衡，但是在更多的时候则不会收敛，因为永远都会发现新的假说、新的想法。

理论很重要，经验也很重要

接下来，我将从经济学的建模转向一个完全不同的领域——经济学的教育。这个领域也可以从认知科学的洞见中受益匪浅。

在这里，我想先提请读者注意，我们要理解某事物，或者说使某事物有意义，有两个途径，即我们可以构建两种不同类型的关联。这两种关联不是完全不同的，不过确实分别位于一个频谱的两端。我把其中一端称为“理论”，另一端称为“经验”。

理论就是包含了蕴涵关系的隐喻。如果在1705年，埃德蒙·哈雷（Edmond Halley）接受了牛顿的引力理论，并应用这个理论分别对在1531年、1607年和1682年的一颗彗星的轨道进行分析，那么就可以得到这样一个蕴涵关系：这颗彗星将于1759年“返回”。在应用牛顿理论时，哈雷实质上是在彗星和牛顿所研究的天体之间建立起了一种关联，而且这种关联的蕴涵关系允许哈雷做出这类预测。我在这里要指出的是，理论是一种“薄”的关联。因为只有在满足一组有限和精确的条件的

情况下，理论才会很好地拟合现象，而且蕴涵关系本身也是有限和精确的。假如正确地拟合了，或者说假如理论给出了一个好的关联，而且自身是内在一致的，那么其蕴涵关系就是可靠的。理论描述的是有限的拟合、有限的蕴涵关系，正是在这个意义上，我们说理论是“薄”的，但是其关联则是强有力的。

经验则与理论不同。假设我是一名经理，被派往韩国开发业务，但是我从来没有去过那里。当我到达韩国后，我想知道我应该如何行事。我不知道与当地人见面时我应该向他们鞠多少次躬，甚至不知道我是不是应该向他们鞠躬。我也不知道进餐时是不是应该把鞋脱掉。我也不清楚，如果我想结束交易，我是不是得等到晚餐结束之后？但是我有许多在日本和中国做生意的经验，所以我可以利用这些经验。在这种情况下，当我到达韩国后，我的脑海中会浮现出许多画面。这也是一种关联，但这种关联有点像梦境，它更加丰富多彩。它还涵盖了更广泛的案例。当给定了某些东西，要预测将会有什么东西随之出现时，经验是非常有启迪价值的。但是，经验与理论相比，要不精确、不准确和不可靠得多。但是无论如何，体现为记忆和情况的图景的经验，是一种“厚”的关联，也是强有力的。它的力量体现在它的覆盖范围和启发性上。我们通常会从人们的对话和故事、小说和戏剧中去寻求这种经验。这就是说，我们会试图吸收别人的经验，使“他们的情况”进入我们的记忆图景，以便供我们日后所用。通过这种方式，我们构建出或想象出一个“梦想”般的世界。在这个世界中，逻辑是不重要的，精确度也是无关紧要的，但启发性和覆盖面却自有其威力。

但是，正如我在前面已经指出过的，这两种类型的关联并不是完全水火不容的。关联是一个频谱，一端是“窄”的“准确性”，另一端是“宽”的“启发性”。

那么，这与经济学研究生教育又有什么关系呢？有很大的关系。教育在很大程度上就是要让受教育者懂得如何联想或建立关联，而且关联的频谱范围也是从一端的狭窄的、精确的理论，到另一端的宽泛的、启发性的和不精确的图景。作为人类，我们同时需要这两种类型的关联。

在经济学中，接受研究生教育的人，必须在一两年内掌握至少20或30个经济学理论模型，即“薄”的关联。这些理论模型包括委托代理模型、世代交迭模型、囚徒困境模型等。这种教育隐含的理念是，待经济学研究生毕业后，这些理论隐喻会在他们的工作中化为有用的关联。我们希望，如假设某个经济学研究生毕业后在世界银行就业，那么只要他

观察到某种情况就能够说出一番“道理”来。比如说，“非洲农业中出现的这个问题部分是一个委托代理问题，某些因素则可以用世代交迭模型来处理，同时还包含了博弈论的成分。所以，我可以建立一个把这三个理论模型结合起来的混合模型，以提出关于这个问题的更深刻的洞见。”这当然很好。好就好在经济学已经认识到，它已经使理论具备了一种循环结构。我们希望，同时也可以合理地预期，一个受过良好教育的经济学研究生会以这些理论为组件，建立起一系列关联。

但是模型并不是我们所应该教的一切。很长一个时期以来，许多研究生院一直倾向于强化理论教学，而不惜牺牲经济史和案例教学。至于学生自己，当然还可以选择，如他们可以去自学经济中的各种经验或细节，但是他们很清楚这种努力对他们毕业后的职业生涯并没有什么助益。1990年，科兰德和克拉默（Klamer）问学生这样一个问题：“对经济的全面了解”对作为一名成功的经济学家的重要性如何？只有3%的人认为非常重要，认为不重要的人却高达68%。在这些学生眼中，更加重要的是“要聪明，即要善于解决问题”和“数学能力突出”。带着这种偏爱理论、鄙视经验的倾向，我们取消了来自历史-经验的更广泛的隐喻，即“厚”的关联。来自历史-经验的更广泛的隐喻，可以让学生将他们的模型置于一个适当的视野中，或者说理论提供了语法，它们则提供了词汇。他们所能提供的丰富的思想和广泛的关联，是理论不可能相比的。

当决策者面对着高度复杂的情况时，如果过于匆忙地应用理论，即一套精确但适用面狭窄的隐喻，就很可能会导致非常危险的后果。波斯尼亚在20世纪90年代中期发生的一系列事件证明了这一点。不妨假设，美国国务院负责波斯尼亚事务的那位高官，毕业于名校政治学研究生院，掌握了许多理论模型，但是却没有多少实际经验。他的反应有可能是，匆匆忙忙地将波斯尼亚硬塞进一个预先构建好的框架内。但是在这种情况下，最好的反应是先等待一下，仔细观察一番。同时，在观察的过程中构建一个随时可以变动的各种图景的集合，利用这些图景可以建立更加丰富的一系列关联。这种“自由关联”源于对历史的研究，而不是死板的理论。如果采取这种思维方式，那么这位官员就会发现：“哦，现在的局势与1908年的波斯尼亚危机有点接近，而且与这个地区处于1831年土耳其统治时的情况相雷同。从另一个角度来看，它的一些因素又类似于1875年的民族冲突，后者最终导致了奥匈帝国的接管……”类似这样的历史图景可能有成千上万，它们都是有用的，可以用来拓展思路、值得考虑。最终，从无数次这样的思考和权衡中，即从大量做梦般的联想中，会涌现出一系列假说或复合图景。只有到了这个阶段之后，

才可以考虑如何应用理论的问题。过早建立理论关联，而不先去考虑一下大量的可能图景，结果将会是灾难性的。在我出生的地方，贝尔法斯特这个城市与另一个非常复杂的情势密切相关，人们经常说："如果你从来没有困惑过，那么你其实什么都不知道。"诚哉斯言！

我之所以说这么多，是要强调学生不仅需要理论，还需要经验和细节。这也就是说，他们需要学习经济史。经济学不是作为理论模型的辅助材料，它自己就是认知意义上的理解的"供应商"。那么，经济学思想史的教学又当如何呢？经济学思想史是经济学教育中另一个有"生存危机"的学科。从认知的角度来看，经济学思想史可以让我们对所做的联想有更加深刻的认识。没有这种认识，联想就可能是"下意识"的，不太可能适合我们要面对的情况。例如，考虑那些在17世纪末或19世纪初来到澳大利亚的英国画家的经历。这些艺术家希望将澳大利亚的树木画入画中，因为他们在英国时就已经学会画树了。他们在英国的艺术学校中学习过，个个训练有素，知道如何画树。但是在澳大利亚，大多数树，典型的如桉树的叶子都比较薄，阳光可以穿过树叶，因此澳大利亚的树看起来与英国的树有很大的区别，它们枝叶更稀疏、更不挡风。结果是，这些来自英国的画家没有画好澳大利亚的树。事实上，澳大利亚本土出生的画家，整整花了一代人的时间，才画出了看上去真正像澳大利亚的树的画。在此之前，来自欧洲的画家会在不知不觉间进行"欧式"联想，并"强加"给澳大利亚。同样地，这个时期的欧洲画家，把非洲原住民画成了"黑皮肤的欧洲人"。我在这里并不是要批评这些艺术家，而只是想提醒读者注意，我们采取的行动其实建立在我们的无意识联想之上。我们必须搞清楚，我们的想象是什么、它们来自哪里。我们必须对它们持怀疑态度。我们需要一种"禅意"，即我们必须先超脱一点，以一个"初生儿"的心智角度观察一番。我们必须意识到，理论不是外生的。恰恰相反，它们是人构建出来的，而且这个"人"在其他情况下也要有计划，那些计划在那些情况下有时是适合的。因此，我们只有掌握了经济思想史，才能对我们在经济学中建立的关联以及它们的出处有充分的意识。

那么，在经济学研究生教育中，我们真正需要的到底是什么？我们当然需要理论。作为理论经济学家，我怎么会说我们不需要理论呢？但是，我们也需要对历史和制度研究提供丰富的图景。我们同时需要两种类型的关联：我们既需要定量的、精确的理论框架，还需要成千上万的基于案例的、生动的图景。只教授理论，就相当于仅仅通过教授内科学和病理学来训练医生，而不让学生参与诊断和实习。无论掌握了多少理论，都只是初学者。在这里，可以想一想通过了"理论考"的学车者。在

我们能够无缝地将理论与生动的图景，即理论与经验整合在一起的时候，我们才可能成为专家。我认为，我们现在教出来的学生当中，许多都只掌握了理论，而没有学会利用图景。我们这样做，是在害他们。

认知真的有那么重要吗

我要求我的经济学家同行们考虑认知的影响，这会不会是个有用但不必要的“奢侈”要求？我不这样认为。鉴于经济学目前的发展水平，我们自然而然地会倾向于那种尽管有价值但显然完全出于想象的一般均衡结果：各项制度都已到位、市场将顺利运行，同时所有激励都是正确的、恰当的。

经济主体不仅有自身的偏好和禀赋，而且有他们自己的理解，即他们会从以往的行动和经验中获得关联和意义。在面对那种“标准化”的、微不足道的经济学问题时，我们可以忽略这一点。但是，在涉及发展和重建的重大问题时，在为创建一门适用于复杂性和界定不明确问题的经济学时，我们就绝对不能忽略这一点了。是的，我们必须认真对待认知。

COMPLEXITY

AND THE

ECONOMY

11

确定性的终结

不确定性是经济世界的主旋律

很多年来，确定论和理性主义思想主导了理论经济学，这让许多经济学家深感不安。如果我们以完美的演绎推理预期采取行动，那么我们的预期就创造了正在试图预测的世界。但如果没有关于他人的预期的知识，任何一个行为主体的预期，从逻辑上看就不可能形成。实际上，经济中存在着根本的不确定性。

在20世纪80年代中期，确定论和理性主义思想主导了理论经济学，但是许多经济学家却对此深感不安。本章是对这种狭隘的“高理性”思维的一个反应，它指出经济中存在着根本的不确定性。我们根据预测或预期采取行动，因此我们的预期创造了正在试图预测的世界。没有关于他人的预期的知识，或者没有关于每个人都知道他人的预期的知识，任何一个行为主体的预期在逻辑上看都是不可能形成的。所有其他人的预期

也是这样。因此，理性预期就成了一个可疑和脆弱的假设，而现实世界则受到了根本的不确定性的影响。这种不确定性意味着，经济行为是建立在主观信念的基础之上的，而主观信念是涌现出来的、共同进化的、不断变化的、相互加强的、相互否定的、会衰变的。我认为，经济学应该承认并接受这样的主观推理，并以之为基本前提。

本章原本是我于1999年在布鲁塞尔举行的“当爱因斯坦遇见马格利特”研讨会上的一次演讲，它里面有一些例子曾经出现在本书其他章节中。本章最初发表于那次会议的文集中。

* * *

20世纪的科学发展史可以说是决定论和确定性一步步丢城失地的历史。许多在20世纪初时仍然像机器一样真实、客观和确定的东西，到了20世纪中叶，已经变得像幻象一样不可预测、主观和不确定了。20世纪初用来定义科学的那些性质，如预测能力、清晰的主客观区别等，都不能再定义科学了。“科学后的科学”已经失去了它的“清白”，但这恰恰说明，“科学后的科学”成长起来了。

经济是确定的吗

首先，经济学是一门科学吗？是的，我相信它是的。经济学肯定是一整套得到了合理论述的知识。然而不同的是，直到最近这几年，经济学仍然保持了它的确定性，或者说它的“清白”没有受到任何损失。所以我们必须问：莫非经济学的研究对象，本质上就是不确定性的和非决定论的？抑或经济学也正在失去其“清白之躯”，也将加入其他“复杂性”科学的行列？

我认为，后者是对的。最近一段时间以来，在经济学中有越来越多的迹象表明，这门科学正在失去其僵硬的决定论立场，长期占支配地位的实证主义思想对它的影响也正在削弱。机械性更少、有机性更多的方法正在步入经济学的舞台。在这个演讲中，我将从我自己的角度来阐明这种确定性的丧失。我将论证，经济中存在着很大的不确定性。我将证明，经济学中以往明确的主体／客体区分，其实是模糊不清的。我将力图阐述清楚，经济不是一台巨大的机器，而是它的行为主体的一个建构物。所有这些，都不是“可怕”的“异常”现象，而是经济的自然属性。如

果我们承认它们，我们就可以将经济学建成一门更加强大的、而不是更加羸弱的科学。

接下来让我从源头说起。经济学的基本思想主要源于18世纪一些伟大思想家的思考，其中特别是来自英国和苏格兰启蒙思想家的思想。1733年，作为启蒙浪潮的其中一个高峰，亚历山大·蒲柏（Alexander Pope）的《人论》总结了启蒙思想的精华：

> 整个自然都是艺术，只是你未曾领悟；
>
> 一切偶然都有方向，只是你没能看清；
>
> 一切不谐，都是你不理解的和谐；
>
> 一切局部的恶，都是整体的善。
>
> 高傲可鄙，只因它不近情理。
>
> 凡存在皆合理，这就是清楚的道理。

在上面这种语境下，“艺术”意味着技巧，即意味着技术或机制。所以，蒲柏说，我们在自然界中看到的一切复杂的奇迹，实际上是一台巨大的机器，一台像他那个时代的机械钟表的机器。所有看上去复杂的、偶然的东西，背后都有自身的运动方向。所有看上去复杂的、无序的东西，都有隐藏的简单性，就像开普勒和牛顿等人看到的行星运动一样。所有对每一个“神的造物”有不利影响的，都以某种无法言说的方式有利于整体的利益。最后，蒲柏引用苏格拉底的名言说：“凡存在皆合理。”

这些不仅仅是蒲柏一个人的观点。在亚当·斯密长大成人的那段岁月里，类似的话语充斥着思想空间。在他的《国富论》中，斯密继承和发扬了这种思想。在这部巨著中，斯密阐明，商人和制造业者、屠夫和面包师等人复杂多样的经营活动背后隐藏着简单的逻辑：经济其实是一门“艺术”，只是它的原理隐而不彰。个人的私利在一只“看不见的手”的引导下，促进了所有人的共同利益。因此，“凡存在皆合理。”两个世纪之后，科学哲学家雅各布·布罗诺夫斯基（Jacob Bronowski）指出，经济学从来没能从18世纪强加于它的致命的“理性结构”中恢复过来。然而，传统经济学继承的不仅仅是斯密的理性结构。在经济学的传统思想的历史深处，还继承了经济就是一种艺术的观念，即经济是一台巨大的机器，只要我们理解了它的各个组成部分，我们就可以预测它的整体行为。当然，在25年前，当我在加州大学伯克利分校修读经济学时，许多

经济学家，包括我自己也曾那样想过，建立经济学的大统一理论是有可能的。从理性人类行为的公理出发，就可以构建出消费者理论；从消费者理论和相应的厂商理论出发，就可以构建出内在一致的微观经济学；而从微观经济学出发，就可以构建出关于经济总体的理论，即宏观经济学。所有这一切将构成一个关于经济的大统一理论。

然而，这种还原主义的宏图伟业，从最底部的组件开始构建整个理论大厦，却有两个难解的问题始终挥之不去：第一，经济依赖于人类，而非依赖于有序的机器组件，因为每个人都有自己的信仰、情感和偏好；第二，经济中的技术是不断进化的。技术会破坏“整洁性”，因为它使经济持续变化。在经济学中，人类行为是利用“经济人”这个工具加以“巧妙处理”的。“经济人”是完全理性的，能够以完全演绎的方式推断出有明确定义的问题的解。至于技术进化，则更难“巧妙处理”，因此不是被忽视，就是被视为外生的。因此，为了得到一个有序的、可预测的理论，经济人，即行为主体，必须只对有明确定义的问题对象进行操作。无论是行为主体，还是要解决的问题，都不得有任何模糊之处。明确定义的问题才有明确的解，而且得到的解将成为下一个聚合层次的理论的构件。

当然，这种方法可以取得一定成效。但是，当问题涉及了不止一个决策者并涉及了任何程度的复杂性时，这种方法就会遇到困难。然后就必须“英勇”地做出大胆的假设，否则良好的定义就不复存在，行为主体和问题就会变得模糊，而不确定的“口袋”就开始膨胀。

“起飞时段选择”的困惑

接下来，我举一个例子，说明上述方法在现代经济学的一个分支——微观经济学的语境下的典型含义。这个例子是我在20世纪70年代后期的产业组织文献中选择出来的。[1]考虑这样一个问题：我们有一个飞机场，24小时都可以供飞机起飞，有许多航空公司，比如说20家航空公司，都在利用这个飞机场，它们必须决定让自己的飞机在哪个时间段起飞，如从拉瓜地亚机场去华盛顿。当然，不同的航空公司有不同的偏好。它们知道自己的偏好，并准备预定适当的起飞时段。所有选择都是一劳永逸的，一经选定就一直适用。在他们真正想要起飞的地方，与不要太接近其他航空公司所选择的这两个起飞时段之间，存在着权衡。在每一个合格的经济问题中，都总是存在权衡问题。所以，给定航空公司

的偏好，他们会选择哪些起飞时间，这就是要解决的经济问题。

将这里的情况说成是确定的，也许有人会觉得有些不对劲，因为它看上去挺复杂、挺不和谐的。但是，在这种不和谐的情况下找到一个和谐的解，正是我想向你们展示的经济学中的现代启蒙方法的精要所在。这种现代方法被称为理性预期。我先将它讲清楚，然后从现实主义的角度去分析它、拆解它，使它露出不确定性的端倪。下面就开始。

在理性预期方法中，首先假设我们知道各航空公司提交选择的顺序。现在不妨先看一下第20家航空公司的选择，想象一下它的推理过程："只要知道前19家航空公司将选择什么，我会知道我自己想选择什么。"因此，无论任何前19家航空公司的选择是什么，我都知道自己将选择哪个起飞时段。对于第20家航空公司来说，这是一个很容易解决的问题。那么，第19家航空公司会怎样选择？第19家航空公司在选择时，知道前18家航空公司的选择，能够计算出自己应该怎么做。给定第20家航空公司，将根据其他18家航空公司和第19家航空公司的选择，来选择一个最佳起飞时段。接下来，第18家航空公司又会怎样选择？没问题。第18家航空公司知道前17家航空公司的选择是什么，知道第19家航空公司将会怎么做，也是可以解决这个最优选择问题的。因为第19家航空公司将会做出自己的最佳选择，而第20家航空公司也将基于第19家航空公司的选择结果，做出自己的最佳选择。复杂吗？这是挺复杂的。但是你可以通过逆向演绎法，或者更准确地说，通过动态规划以相反的顺序搞清楚这里的逻辑，并推断出所有20家航空公司将如何做出自己的最优选择。

需要注意的是这个推理过程的性质。问题是有明确定义的。这是一个序贯决策问题，而且假设所有航空公司都运用逻辑上的逆向演绎推理法，得到的解在数学意义上是精确的、"干净"的。因此，这个经济问题变成了一个数学问题。所有这类问题都会变成数学问题，从而使经济学变成了数学。这类问题通常还有另一个属性，那就是个人行为对整体是有好处的，也就是说，部分的恶是普遍的善。尽管在现在这个例子中，情况不是这样，但是在经济学中，这是一个一般属性。

然而，这个解还带有很多的"附带条件"。各航空公司必须准确地知道自己的偏好，不仅如此，它们还必须准确地知道所有其他航空公司的偏好。此外，它们还必须知道，每一个其他航空公司都准确知道每个其他航空公司的偏好。它们还必须知道，每家航空公司都知道，每家航空公司都知道每家其他航空公司的偏好……这是一个无限的递归。此外，

每家航空公司还必须足够理性，以求出那个最优解。更进一步，每家航空公司还必须确信，每一家其他航空公司都是理性的，并且都会运用完全理性来制定解决方案。此外，每家航空公司都必须知道存在一个无限的递归，知道每一家其他航空公司都正在使用这种理性的方法来解决这个问题，因为这些航空公司当中只要有一家“搞砸”了，就会使每一家其他航空公司的解全都乱成一团。在使用这种逆向演绎法求解时，每家航空公司的最优解都必须是唯一的。在这个“需求网络”中，任何一个链条一旦中断，解也就不复存在了。对此，我的评语与那位法国元帅康罗贝尔（Canrobert）在比利时留下的评语一模一样：“这确实极为壮观，但是这不是在作战！”

这种类型的多行为主体选择问题在经济学中是非常普遍的。所以，我们必须认真对待这种求解方法。假设我们是第三家航空公司，如果我们不能确定第17家航空公司将会怎么做，那么我们应该怎么做？作为第三家航空公司，我们也许会说：我不相信第17家航空公司的人都有那么聪明，我不知道他们是否会运用上述完全理性的方法解决这个问题。如果他们不这样做，那么我不知道我作为这个序贯决策问题的第三个决策者的最佳选择是什么。这种不确定性足以打乱整个情况。而且更加糟糕的是，第三家航空公司可能会将这种不确定性“传染”给其他航空公司，于是它们就可能无法再依赖第三家航空公司和第17家航空公司。这样一来，整个解也就开始土崩瓦解了。得到的解是航空公司的预期或对其他航空公司将要做什么的预测的函数。因此问题的要害是，如果我是一个有“代表性”的航空公司，我要找到的是我的预期应该是什么，这也就是说，我正在试图预测一个由我自己和所有人的预期创造的世界。这里存在一个自我指涉的循环。每家航空公司试图预测的结果，取决于它和其他航空公司可能形成的预测。换句话说，这些预测正在形成这些预测试图预测的世界。除非有某种协调装置，使得各航空公司可以从逻辑上确定其他航空公司的预测，如上面这种“折磨人的”逆向演绎推理方法，没有其他方法可以确定其预测。

因此，这里存在一种逻辑不确定性。这就是说，在经济中，人们正在创造一个从他们的预测中形成的世界，但是如果他们试图以一种完美的演绎推理的方式形成这些预期，那么他们就会进入一个自我指涉的循环。标准的经济学思想存在一个逻辑漏洞。我们的预测共同创造了我们的预测正在尝试预测的世界。如果不知道别人如何确定他们的预测，那么我的预测就是不确定的。当然，在经济学中，在某些情况下，很显然每个人都知道自己该做什么。在这种情况下，像上面提到的那种方法是

有效的。但是除了这类情况之外，我们在所有其他情况下都会面临这个根本的不确定性问题。经济中的行为主体生活在一个马格利特式的世界里。当我们的想法和偏好共同创造了它们试图预测的世界时，自我指涉就会使得问题变得不确定。我们可以将经济的主体，即形成经济的行为主体与客体经济完全区分开来，这种想法是有问题的。不确定性的“口袋”在经济中随处可见。现代形式的经济决定论注定要失败。

“资产定价”的困惑

在这里，有两个问题是我们要追问的。第一个问题是：这一点很重要吗？也许所有这一切都只发生在一个零测度集上，也许这种困难仅仅局限于经济学中的一些琐碎的、无关紧要的例子中。第二个问题是：如果真的很重要，那么我们应该怎么办？

接下来，我就来回答这两个问题。我将举出一个来自资本市场领域的例子，即资产定价理论。我可以先告诉大家，这个理论与前面解释航空公司的选择理论遇到了同样的麻烦。

唯一的区别是，资产定价理论是一个非常重要的理论。

1991年，我在香港受聘担任花旗银行的顾问。该银行试图开发一个复杂的神经网络模型，用以预测外汇市场的价格变化。作为经济学家，我的初步反应是对此表示怀疑。我相信标准的有效市场理论，它的其中的一个结论是，没有办法预测金融市场。但是很快地，我就发现外汇市场上的交易者不同意这种看法。至少看在能帮助他们赚钱的事实上，他们认为他们可以预测价格走势。不过，我还是先得快速地概述一下标准理论。标准的有效市场理论说，所有进入金融市场的信息都会被投机者和投资者利用，而且对未来价格变化有任何预示的所有信息都会被“用尽”。换句话说，通过一个与上面那个航空公司的例子非常类似的推导过程，每个投资者对每只股票的出价都会定格到一个唯一的水平上，它只取决于当前可用的信息。根据这种观点，使用过去的价格模式来预测未来价格，即技术交易方法是不可能带来更多的利润的。不然的话，过去的价格内含的信息可以用于获得更多利润，并且根据假设，投资者已经将所有信息都折现为当前价格了。因此，标准理论认为，投资者利用所有可用的信息形成预期。这些预期将会确定股票的价格，而平均的价格又将支持同样这些预期。这又一次说明，理性预期是存在的。因此，在这种情况下没有办法赚到钱，市场是有效率的。

但是，交易者则认为市场是可以预测的。他们认为，可以从过去的价格中发现某种有助于预测的模式，即他们相信技术交易。他们认为市场是拟人性的，也有自己的“心理”和“动机”，如“市场昨天一度很紧张，但是它对坏消息没有太多过激反应，最终稳下来了”。经济学家则对此持怀疑态度。我记得，我曾经听一位著名的经济学家说过：“如果技术交易可以赚钱，那么就会出现很多极其富有的公司和银行。”这句话着实令我迷惑不解。因为在现实中，确实有很多公司和银行通过各种形式的技术交易，变得越来越富有了。

当然，标准理论是很成功的。它有它自己的逻辑。这个逻辑是完整的，而且具有所需要的一切属性，如数学上的唯一性。但是，标准理论必须面对一些不明原因的现象所带来的挑战，它称这些经验现象为“异常现象”。“异常现象”一词的基本含义是，这些现象是有“问题”的，因为它们不符合理论。据说，“异常现象”并不能说明理论有问题，尽管它不能解释这些现象。因此根据标准理论，即使1987年10月的股票市场和其他金融市场损失了23%以上的价值，也只能被称为一种“修正”。但是，1987年10月并没有任何消息要对这次崩溃负责。另一个异常现象是“泡沫”，如著名的荷兰郁金香泡沫。“泡沫”指价格保持在极高的水平上，而且没有任何明显的理由。此外，市场交易量也比理论预测的高出了几个数量级。已经有好几位经济学家，如布洛克、拉克尼肖克（Lakonishok）和勒巴伦等证明，技术交易在统计上确实是有利可图的。另一个谜题是所谓的GARCH行为，它指的是股票价格的高波动期与静止期随机地交替出现。

总之，资产市场上至少存在6个重要的统计上的异常现象，是标准理论无法给出解释的。这也导致了许多更现代的、更巧妙的新理论，有些运用了特殊的行为观察结果，有些则给出了更加复杂的理论模型。

现在我来阐明，对于资产市场，标准理论是如何崩溃，并导致了不确定的“口袋”，就像前面的航空公司的例子一样。假设投资者可以将他们的一部分资金投入单一的一只股票，它每一期，比如说一天或一年支付一定股息，但股息的具体数额投资者无法完美预测。投资者购买股票，是为了获得股票的股息和任何资本增值（未来的价格），因此他们面临着如何预测未来的股息和价格的问题。为了得到标准理论的解，我们假设投资者是同质性的，即全都是相同的“克隆体”，他们对期末股利的预测、对股票未来价格的预测全都相同。预测的平均结果是无偏的，因此这是理性预期。只要稍作经济学推理，就可以证明，今天的价格等

于明天的价格加上股息，再经适当贴现和加权后的共同预期。这样在每一期都会产生一系列方程。然后利用条件期望代数方法，我们就可以解出这些方程，求得以当前信息为条件的对未来价格的预期，最终得出作为预期的未来股息函数其当前价格的表达式。问题解决了。但是，这只是给定了“假设投资者是同质性的，即全都是相同的‘克隆体’，他们在期末对股利的预测、对股票未来价格的预测全都相同”的条件下的解。如果我们不做这个假设会怎样？如果我们假设投资者不是完全相同的，又会如何？

接下来就让我们来看看，假设我们的投资者行为主体不是“克隆体”，即不是同质性的，运用与上面相同的推理方法会得到什么。这里需要注意的是，标准理论对完全相同的“信息”的要求，不仅意味着每个人都会观察到完全相同的数据，而且意味着他们对数据的解释也完全相同。但是，请想象你自己在一个真正的市场环境下的情况，如在香港外汇市场中，信息包括了过去的价格和交易量、新西兰中央银行或新加坡银行或中国人民银行采取的行动、谣言、CNN的评论、新闻、你的朋友在做什么、他们通过电话告诉你什么、某人的姨妈认为市场将发生什么……不一而足。所有这些都包含了实际信息，而且我们可以合理地假设，即使每个人都拥有访问所有这些信息的相同的权利，他们也会将该信息视为“罗夏克墨迹测验”中的墨迹，并对其进行不同的解释。即使我们假设解释这些信息的人拥有无比强大的智能，也可以说是“无限聪明的”，而且他们受过系统的统计学训练，他们也仍然会对数据给出不同的解释，因为解释相同的数据有很多种不同的方法。

因此，不存在单一的预期模型。每个个体投资者仍然可以得出个人对股息的预测。但是未来的价格则由投资者和其他投资者对股息和下一期价格的个人预测所决定。而且，个人投资者无法预先知道其他人的预测，借用凯恩斯的话来说，即不能算出预期的“平均意见的平均意见是什么”。那样做的话，会导致逻辑上的递归：“我认为他们可能会认为，而且同时还意识到我会认为他们会这样认为。”除非我们假设投资者是完全相同的，不然的话，我们的行为主体就还在尝试预测作为其他行为主体的预测函数的结果，即未来价格。这里的问题与前面举的航空公司的例子中的问题是一样的，都是演绎无法闭合。从而，预期变成了不确定性的，理论失败了。

而且更加糟糕的是，预期还会变得不稳定。请想象一下，假设有几个投资者认为市场上的价格会上涨。如果我相信这一点，并且我还相信

别人也相信这一点，那么我就会向上修正我的预期。但是，后来我又收到了一些负面的传言，于是我会重新评估，而且意识到别人也可能会重新评估，同时意识到他们也意识到别人，即我可能会进一步重新评估。这样一来，无论交易是否会发生，预期就会变成一种不可捉摸的转瞬即变的东西，于是价格就不断上涨或下跌。从而，预测变得不稳定了。这是价格泡沫开始出现的方式。如果人们预期价格上涨，他们会预测其他人也会预测价格会上涨，所以他们会买入，从而启动了一个泡沫。当人们观察到价格上涨后，他们对价格上涨的预期就实现了，因此价格可能继续上涨。这就是说，泡沫是自我实现的。

类似的逻辑同样适用于对“价格底”和“价格顶”现象的分析。举例来说，如果当前价格是894，同时许多投资者都认为在900的位置有阻力，那就会成为一个“价格天花板”，即一个“顶”。当价格达到这个上限时，它有一定概率回落，或者也可能“突破”这个“天花板”。我听到“价格底”和“价格顶”这样的叫法时，第一反应是怀疑，不过很快我就开始意识到，许多投资者之所以在900的位置卖出股票，只是因为它是一个整数。但是，这将使价格触及900后会出现回落这种预期成为一个自我实现的预期。“价格底”和“价格顶”是作为部分自我实现的预期而涌现出来的，这就是说，因为在这些位置买卖股票比较方便，所以价格维持住了。现在，我们已经偏离同质的理性预期相当远的距离了。在更加现实的假设下，交易者可能对不同的信息给出不同的解释，从而使预期变得不确定和不稳定。而且，它们很可能会相互自我实现。

上面说的这些可以总结如下：只要我们仔细观察资本市场理论这个重要的经济学分支，我们就可以看到与航空公司的例子中同样的不确定性。行为主体需要对结果形成预期，而结果是所有这些预期的函数。只要假设对“信息”的解释存在合理的异质性，演绎推理就无法闭合。因此预期的形成是不确定的。

现实世界是这样的

但是我们来看一看现实世界！在现实世界的每一个市场中，人们每一天都要形成预期。那么他们如何做到这一点呢？如果他们无法以演绎推理的方式做到这一点，那么我们又应该怎样对他们在这方面的行为建模呢？1988年，约翰·霍兰德和我决定，我们将通过一个基于计算机的人工股票市场来研究这个问题。在这个人工股票市场中，“小小的”行为主

体，实际上是一些人工智能计算机程序，拥有某种手段来完成这里所需的推理。这是最早的基于行为主体的人工市场之一。后来，我们又邀请了物理学家理查德·帕尔默、金融专家保罗·泰勒，还有金融学家布莱克·勒巴伦加入我们的研究项目。当我们的研究刚启动时，著名的计算机科学家约翰·霍兰德，虽然是遗传算法的设计者，但是他只会用BASIC编程，我也只会用BASIC编程。幸亏，帕尔默是一个高明的程序员，所以我们的进展非常迅速。我们先在麦金塔计算机上实现了基于计算机的人工股票市场，后来又移植到了NeXT计算机上，并用它完成了一系列实验。

我们没有往这个人工股票市场输入任何来自实际股票市场的数据。这是一个存在于计算机内的人工世界。它的人工行为主体是一些人工投资者，在彼此之间买卖一只“股票”。计算机可以显示股票的价格和股息、谁在买卖、谁赚了钱、谁亏了钱、谁进入了市场、谁留在市场外面，等等。价格在计算机内通过出价和要价形成。另外还有一个小程序的“专家”，它设定市场出清价格，就像在实际股票市场上一样。

这个模型试图解决的问题是：如果行为主体不能以演绎推理的方式形成他们的预期，那么他们又将怎样才能形成他们的预期呢？我们决定遵循关于真实人类在这种情况下会如何行事的现代认知理论。因此，我们允许我们的人工行为主体查看股票价格变动的最近历史，建立多个用来预测的假说或模型，并不断地对它们进行测试。这些假说中的每一个都具有与之相关联的预测。在任何阶段，每个行为主体使用其最准确的假说，并相应地买入或卖出股票。在我们的人工市场中，行为主体通过两种方式进行学习：第一，他们能够学会判断他们建立的预测假说当中哪一个更加准确；第二，他们会不断地舍弃那些不起作用的预测假说，并利用遗传算法产生新的预测假说将它们替换掉。因此，他们一直会学习识别他们共同创造的模式，这反过来会共同创造股票价格的新模式，对此新模式他们又可以形成新的假说。提出假说、测试它们并偶尔替换它们的这样一种行为，就是我们所称的归纳。因此，我们的人工市场的行为主体使用的是归纳理性。这无疑是一种更加现实的行为方式。

这一切看上去还不错吧。但是，现在还有一个关键问题，那就是我们的人工市场是会收敛到标准经济学理论所说的理性预期均衡，还是会表现出某种截然不同的行为呢？我们惊讶地观察到，在我们的人工市场中，这两种“市场体制”都出现了。第一种体制，我们称之为“理性预期体制”，当我们让人工市场的行为主体用非常接近理性预期的预测假说开始

他们的交易时出现了。我们在图表上描绘出了预测假设的所有参数，结果发现在这种情况下，随着时间的推移，它们被引入了理性预期解的轨道。所有预测假说围绕理性预期均衡点形成了一个“毛带”（fuzz），那是因为它们偶尔会偏离理性预期以测试不同的想法。为什么理性预期会占上风？理由很简单。如果预测整体上接近理性预期，那么价格序列将接近理性预期，这样非理性预期预测将被“证伪”。所以在这种情况下，标准理论得到了验证。

但是还有第二种体制，我们将它称为“复杂体制”。在更广泛的情况下，都是这种体制占上风。我们发现，如果我们让人工市场的行为主体用有些偏离理性预期的预测假说开始，或者如果我们允许他们用比以前稍快的速度提出假说，那么市场的行为就会完全改变。在这种情况下，市场中涌现出了相互加强的预测子集。不妨设想一下，如果我们的人工市场中有100个行为主体，每个行为主体都使用60种不同的预测公式，那么总共就会有大约6 000种预测器。在这些涌现出来的预测器中，有一些是相互加强的，有一些是相互否定的。假设有许多预测器的预测是，股票价格将随时间推移而周期性地上下波动，那么这些预测器就是相互否定的，因为它们将导致行为主体在周期底部买入，并在周期顶部卖出，相互抵销利润，最终会从预测器“种群”中消失。但是，如果出现了这样一个预测器子集，它们的预测是“如果在过去三期内价格是上涨的，那么下一期的价格也会上涨”，而且这样的预测器足够多，那么就会导致行为主体不断买入，这样平均来说会导致价格上升，从而反过来强化这样的子群体。在这种情况下，这类预测器子集将会“突然走红”，并且嵌入到预测器的种群中。在我们的人工市场的“复杂体制”中，确实发生了这样的事情，从而使得该市场的行为更加丰富多彩。这种复杂体制的另一种表现是，我们的人工交易者发现了有效的技术交易形式。他们能够基于过去的价格模式来进行预测，并取得了成功。因此，我们的人工股票市场涌现出了技术交易。而且，这种拥有相互强化元素的子集的涌现，不禁令人想起生命的起源。生命起源的关键就在于，正确组合的RNA亚群涌现出来并可以“相互加强”。这不能不令人惊叹。

复杂体制的另一个涌现性质是GARCH行为，即在股票价格的若干个高波动期与若干个“静止期”交替出现。这也是标准资产定价理论无法解释的一个现象。那么，GARCH行为又是如何成为一个涌现性质的呢？在我们的人工股票市场中，经常有一些投资者发现，某种新的方法能在市场上获得更好的结果。即使改变幅度非常小，这些投资者还是会改变他们的买卖行为，并进而导致市场发生变化，这反过来又导致其他投资者

也改变自己的行为……由此，变化的“雪崩”将席卷整个市场。这就出现了变化触发进一步变化的高波动时期，以及随后出现的几乎没有变化且很少需要改变的静止时期。这就是GARCH行为。

现在，让我总结一下。我们在人工股票市场中发现的是，在我们的投资者从非常接近于理性预期解的位置开始的情况下，这种“解”占优势。但这只反映了参数空间的一个很小的集合。除此之外，复杂体制将占优势。在复杂体制下，会涌现出自我强化的信念和自我增强的“变化雪崩”，然后涌现出更丰富的“解”或行为集，它们符合实际的市场行为。这是一个更一般的理论，理性预期理论是它的一种特殊情况。

源于启蒙运动时的标准经济学观点认为，经济是一个客体对象。它也许是复杂的，但还是可以机械地予以处理。主体与客体，即行为主体与他们所处的经济，是可以被“整齐”地区分开来的。但是，我在这里给出的观点则完全不同。我认为，经济本身就是从我们的主观信念中涌现出来的。这些主观信念聚合起来，就构成了微观经济。它们使金融市场形成了作为金融市场的特性。它们引导资本流动，支配策略性行为和谈判。它们是经济的DNA。这些主观信念是先验不确定的，或者说是事先不可能通过演绎方法确定的。它们共同进化、生成、衰亡、变化、相互加强或相互否定。主体和客体也不能整齐地区分开来。因此，对于经济表现出来的行为，我们只能描述为有机的，而不能描述为机械的。经济不是一个有序的、巨大的机器，它是有机的。在所有层次上，它都包含着不确定性。它涌现于主观性，并将回到主观性上。

COMPLEXITY AND THE ECONOMY

结语

复杂的经济需要复杂经济学

在经济系统中，人类行为主体能够通过思考自己将要采取的行为可能带来怎样的结果，对自己的策略和预测进行调整，因此，完美的演绎推理就失效了。源于收益递增的正反馈是经济系统的重要特征。复杂经济学不是标准经济学理论的附属物，而是一个更一般的、超越了均衡观念的理论。

本章总结了我在1999年对“复杂性与经济”这个问题的思考。它是本书第1章的前身，而且两者有所重叠。尽管如此，我还是决定将它收录进来，因为它第一次引入了“复杂经济学”这个术语。本文最初发表在1999年的《科学》杂志上。

* * *

一切关于复杂性的研究，针对的都是多元素系统，这些元素要适应或对它们自身创造的模式做出反应。这里所说的元素可以是元胞自动机中的元胞，它们会对邻近的元胞的状态做出反应；也可以是自旋玻璃中的离子，它们对局部磁矩做出反应；或者是免疫系统中的细胞，它们对B细胞和T细胞的浓度做出反应。元素也好，它们所要适应的模式也好，都因具体的情境而异。但是不变的是，元素必须适应它们创造的世界，即它们共同创造的总体模式。通过调整和变化的过程，时间因素也自然而然地进入系统：当元素做出反应时，总体模式发生变化；随着总体模式的变化，元素又得重新做出反应。除非达到某种渐近状态或均衡，一个复杂系统必定是处于不断进化和展开中的系统。

复杂经济学是一种超越均衡的经济理论

经济自然是复杂系统。经济行为主体，无论是银行、消费者、厂商，还是投资者，都要不断地调整各自的市场行为、买卖决策、定价策略和对未来的预测，以适应这些市场行为、买卖决策、定价策略和对未来的预测共同创造的情境。但是，与经济系统中的元素不同的是，旋转玻璃中的离子总是以一种简单的方式对局部磁场做出反应。而经济这个系统中的元素，即人类行为主体，能够通过考虑自己所采取的行为可能带来的结果，对自己的策略和预测进行调整来做出反应。这就给经济学研究又增加了一层复杂性，而在自然科学中是没有这层复杂性的。

面对经济这个复杂系统，传统的经济学理论的选择是，不去研究它的行为主体创造的模式的展开，而是简化其问题以便于求出解析解。因此，传统经济学要问的问题是，什么样的行为元素，即行动、策略和预期，是与这些行为元素共同创造的总体模式相一致的？例如，一般均衡理论问的是，在不会产生任何导致改变激励的情况下，生产出来和消费掉的商品的价格和数量，与经济中各市场的价格和数量的总体模式是否一致。博弈论问的是，在给定的策略情境下，什么样的行动、选择或分配，与其他行为主体的行动、选择或分配相一致或是最优的。理性预期经济学问的是，什么样的预测或预期与这些预测和预期共同创造的结果相一致，或能够被这些结果所验证。因此，传统经济学研究的是各种“一致的模式”，即行为均衡中的模式，它们不会引发进一步的反应。不过现在，圣塔菲研究所、斯坦福大学、麻省理工学院、芝加哥大学和其他一些研究机构和大学的经济学家正在突破这种均衡方法，转而研究行动、策略或预期怎样对它们创造的、可能内生地改变的总体模式做出反应。

复杂经济学就是他们这种努力的结果。复杂经济学不是标准的经济学理论的附属物，而是一个更一般的、超越了均衡观念的理论。

一只“轻推的手”

如果我们描述的这类系统，还包含了以正反馈形式表现出来的非线性特征，那么它们就会变得特别有趣。在经济学中，正反馈源于收益递增。为了确保能够得到唯一的、可预测的均衡，标准经济学理论通常需要假设收益递减：如果一个厂商在市场上遥遥领先，那么它会遇到更高的成本或某种其他形式的负反馈，结果市场份额将稳定在某个可以预测的唯一的均衡水平上。然而，当我们允许正反馈或收益递增时，就会产生完全不同的结果。例如，以几年前的在线服务市场为例。在这个市场上，主要是三家公司在竞争：Prodigy，Compuserve和America Online（美国在线）。当一家公司的会员基数得到扩大之后，它就可以提供更广泛的服务项目，让更多的会员分享专门的服务和定制的聊天室，以满足自己的爱好。这也就是说，扩大会员基数是收益递增的。Prodigy是最早进入这个市场的企业，但是后来者美国在线却抓住机会，通过适当的策略获得远超同侪的无可比拟的优势。到今天，它已经占据主导地位了。如果情况稍有不同，那么它的竞争对手就可能已经占据了市场。请注意这里呈现在我们面前的一些性质：存在多种潜在解，实际达到的结果是不可预见的，经常会被锁定，结果不一定是经济上最有效的，结果取决于所采取的历史路径，尽管各公司在开始时可能是平等的，但是结果却是不对称的。在存在类似的正反馈的非线性物理学中，这些性质不难找到各自的对应物。经济学家所说的多重均衡、不可预测性、锁定、无效率、历史路径依赖性和不对称性，大体对应于物理学家所称的多重亚稳定状态、不可预测性、相位或模式锁定、高能基态、非遍历性和对称性破坏。

在经济学中，收益递增问题已经被讨论了相当长的时间。早在100多年前，阿尔弗雷德·马歇尔（Alfred Marshell）就曾经指出，如果一个厂商能够通过扩大自己的市场份额来占据优势，那么“任何率先获得良好开端的厂商都可以获得垄断地位”。但是，传统的静态均衡方法无法克服不确定性带来的障碍：在存在多重均衡的情况下，怎样才能达到其中的某个均衡呢？面向过程的复杂性方法提出了解决这个问题的一个途径。在实际经济中，各种小的随机事件时时刻刻都在发生。例如，在前面举的在线服务的例子中，就会发生诸如随机接口改进、新产品推出、好口

碑、用户推荐这类的事件。

随着时间的推移，收益递增会放大这种事件对随机选择结果的累积效应。因此，在研究经济学中的收益递增问题时，最好将它视为具有随机事件和自然正反馈的动态过程，即非线性随机过程。这种从静态视角到过程导向的转变，在复杂性研究中是极为常见的。现在，收益递增问题在市场分配理论、国际贸易理论、技术选择进化理论、经济地理学，以及对贫困化和族群隔离模式的研究中，都已经得到了大量的研究。在所有这些研究领域，研究者们得到了一个共同发现，那就是经济结构可能因某些小事件固化并被锁定。这个共同发现已经开始改变政策，使决策者意识到政府施政必须避免两个极端：强制实现预期的结果，或者极端放任自由，而应该设法将系统“轻轻地推到”能够生长和自然涌现出结构的那个方向上去。这不是一只“沉重的手”，也不是一只“看不见的手”，而是一只“轻推的手”。

一旦我们采纳了复杂性视角，强调结构的形成而不是强调给定结构的存在，那么经济学中预测问题就会变得完全不同。传统的常规经济学方法试图回答的是，在某个特定的问题中，假设预测模型或预期模型是由所有行为主体给出并共享的，那么什么样的预测模型是与该预测模型部分生成的实际时间序列是一致的，这一般而言是可以得到验证的。这种“理性预期”方法在一定范围内是有效的。但是，它必须假设行为主体能够预先以某种方式推导出什么模型是有效的，并且每个人都知道并使用这个模型，即共同知识假设。关键是，当预测模型不那么“显而易见”，并且对他人的预期不知情的行为主体必须独立成形的时候，又会怎么样呢？

“归纳理性”的胜利

作为例子，考虑我提出的“爱尔法鲁酒吧”问题。假设有100个人，每个人每个星期都必须独立地做出决定，要不要去他们最喜爱的酒吧，即圣塔菲的爱尔法鲁酒吧去消遣。规则是：如果某个人预测同一天会有超过60个人去酒吧，那么他或她为了避免拥挤嘈杂就选择留在家里；如果预测不到60个人会去，那么他就选择去。我们想知道的是，这些酒吧客人怎么预测每个星期将会出现在酒吧的人数，以及由此导致的数字序列的动力学。请注意，这个问题有两个特点。第一，我们的行为主体很快就会意识到，对有多少人会去酒吧的预测取决于其他人对多少人会去

的预测，因为这决定了他们的“到场率”。但是，其他人的预测又取决于他们对其他人预测的预测。运用这种演绎推理时，存在着一个无限的回归。没有任何一个“正确”预期模型可以被假定为共同知识，而且从行为主体个人的角度来看，这个问题无法明确界定。不仅仅这个问题是这样，绝大多数预期问题都是这样。第二，更加令人讨厌的是，任何一个“共同预期”都会被打破或落空。如果所有人都运用预测只有少数人会去的预期模型，那么所有人都会去，从而证伪了这种模型。同样地，如果所有人都相信大多数人会去，那么没有人会去，这样会使这种信念归于无效。因此，各个人的预期将被迫变得不同。

1993年，我针对这种情况构建了一个模型。假设行为主体在决定去不去酒吧时，他们会采取一种归纳推理的方法，即把他们视为统计学家，每个人都从一系列主观选定的预期模型或预测假说开始“工作”。每个星期，他们根据他们目前拥有的最准确的模型来采取行动，我将这些模型称为他们的活跃预测器。因此，行为主体的各种信念或假说要在它们自己创造的“生态”中相互竞争。

计算机模拟结果表明（见图12-1），平均到场人数很快就会收敛到60这个数值上。各预测器自组织成了一个均衡的生态。在所有活跃的预测中，平均来说有40%预测到场人数会多于60人，有60%预测到场人数将少于60人。这种涌现出来的生态从本质上看是有机的，因为虽然活跃预测器组成的“种群”按60：40的平均比率划分为两个“子种群”，但是每个“子种群”的成员永远都在不断地变化。那么，预测器为什么会自组织成平均到场人数为60人、所有活跃预测器按60：40的比率划分这种状态呢？为了回答这个问题，不妨假设70%的预测器在相当长的时间内都预测出场人数多于60人，那么平均来说将只有30个人会去酒吧。但是，这个结果验证了那些预测到场人数接近30人的预测器，从而会将生态恢复到均衡水平。40%：60%这个“天然”组合成了一个涌现出来的结构。“爱尔法鲁酒吧”问题代表了一个具有复杂动力学的微型预期经济。

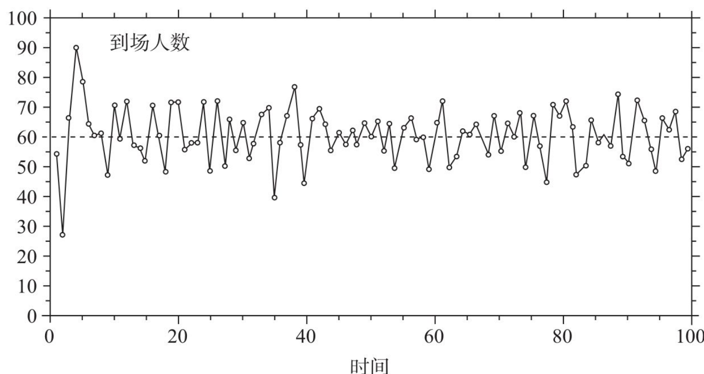


图12-1　该酒吧前100周的到场人数

经济世界怎能少了“泡沫”和“崩溃”

这种思想的一个重要应用是在金融市场领域。关于金融市场的标准理论假设了理性预期，即行为主体会采用统一的预测模型，在平均水平上，模型的有效性可以由其预测的价格得到验证。这个理论在只存在一阶预期的情况下确实很有效，但是它不能解释实际市场上出现的大量“异常现象”。例如，意外的价格泡沫和崩溃、价格变动的高波动期和低波动期的随机交替出现，以及基于近期历史价格模式的技术交易方法的大量使用。

霍兰德、勒巴伦、帕尔默、泰勒和我建构了一个模型，放松了理性预期假设。像在“爱尔法鲁酒吧”问题中一样，我们的假设是，投资者无法直接提出或通过演绎推理的方法推定预期，而是必须自己去发现它们。在我们基于计算机的人工股票市场中，行为主体不断地创造并使用多个关于未来市场价格和股息的假说或预期模型，这些模型都是个人化的、主观的。当然，我们这个人工股票市场中的“投资者”，实际上是一些人工智能计算机程序，他们可以生成和舍弃预期假说，并根据他们目前最准确的假说进行出价和要价。股票价格产生于他们的出价和要价，因此最终形成于行为主体的预期。这也就是说，这个基于计算机的市场是一个自足的人工金融世界。像“爱尔法鲁酒吧”一样，它也是一个小型生态。在这个生态中，各种预期在一个由这些预期自身创造的世界中相互竞争。

在这个基于计算机的人工市场中，我们发现了两种“相”或“体制”。如果设定参数使我们的人工行为主体只能缓慢更新他们的假说，那么预

期的多样性会迅速坍塌成为同质的理性预期。原因是，如果大多数投资者的预期都接近理性预期，那么由此而得到的结果，即价格将验证这种预期，同时预期模型“种群”出现的偏差或突变预测都将被证明是不准确的。在这种特殊情况下，标准的金融理论仍然成立。但是，如果假说的更新率稍有增加，那么市场就会出现“相变”，转入“复杂体制”，并呈现出我们在真实市场中观察到的那些“异常现象”，从而开启了一个丰富多彩的“心理世界”。不同信念，不会随着时间的流逝而收敛。在由各种假说组成的“种群”中，像“如果市场趋于上升，那么预测价格会上涨1%”这样的预期规则一旦随机出现，就会相互加强，因为只要有足够多的投资者根据这类预期规则采取行动，那么价格确实就会上涨。因此，相互加强的预期的子群体出现后，行为主体就会在这些预期上“投下赌注”，这样技术交易就出现了，从而导致偶尔的价格泡沫和崩溃。在我们的人工模拟市场中，也出现了在价格波动的高峰期之后，低波动期随之而来的现象。这是因为，如果一些投资者发现了新的可以带来利润的假说，他们就会使市场出现些微的变化，进而导致其他投资者也改变他们的预期。因此，所有尺度上的信念变化都会像雪崩一样造成市场波动，导致市场上高波动期和低波动期交替出现。实际的金融市场也呈现这些现象，因此我们不妨推测，实际的金融市场是属于这种复杂的体制的。

经济是依赖于过程的、有机的和永远进化的

经济学已经花费了整整两个世纪的时间研究均衡，即不要求进行进一步行为调整的静态模式。现在，经济学家终于开始研究经济中的结构涌现和模式展开了。复杂经济学不是静态的传统经济学理论的暂时附属物，而是更加一般的、超越了均衡层面的理论。这种方法的活力在经济学的所有领域都能感觉到：博弈论、货币和金融理论、经济学习、经济史、贸易网络的进化、经济的稳定性，以及政治经济学。它有助于我们了解和应对诸如市场不稳定性、垄断的出现和贫困的持续性等众多现象和问题。它告诉我们，政策要想成功，只能通过影响经济结构形成的自然过程，而不能通过强行实现某种静态结果。

当然，即使我们超越均衡框架去观察，经济模式有时也会落入标准经济学的简单的、同一的均衡当中。然而，更加常见的情况必然是，经济模式时刻处于变化中，总是为我们呈现新异行为和涌现现象。因此，复杂经济学认为，经济不是确定的、可预测的、机械的，而是依赖于过程的、有机的、永远在进化的。

未来的经济学原则

表14-1是我于1979年11月5日写在我的研究笔记上的，现在我把它收录在本书中。这不仅是因为，在过去很多年里，这张表被引用了很多次，而且还因为本书收录的这些文章的思想源头就在这张表中。

1979年，我在奥地利国际应用系统分析研究所工作。我阅读了很多生物学论著，其中特别是雅克·莫诺（Jacques Monod）和弗朗索瓦·雅各布（Francois Jacob）的工作成果，以及布鲁塞尔和斯图加特的几个研究小组关于自组织系统的研究成果。以这些思想为基础，加上我自己的思考，我逐渐认识到，未来的经济学与传统的经济学必定会大异其趣。未来的经济学是基于不同的原则的。我把我当时想到的“未来的经济学原则”写在了我的笔记本上，结果就是这张表。它还包括了一些关于人口统计学的想法，因为那是我当时最感兴趣的领域之一。

表**14-1**　新旧经济学的对比

旧经济学	新经济学
收益递减	报酬递增将大行其道
以边际主义为基础（新古典）	其他可能的原则（例如会计原则）

续前表

旧经济学	新经济学
最大化原则（利润动机）	秩序原则
偏好给定	偏好的形成将成为核心研究议题
个体自私	个体不一定自私
社会只是背景	制度走上前台，成为各种可能性、秩序和结构的主要决定者之一
技术给定或根据经济来选定	技术从一开始就是可变的，然后倾向于固定
本质上是决定论的、可预测的	非决定论的。因为振荡和奇异吸引子（以及类似的东西）不可预测
建立在19世纪的物理学理论的基础之上（均衡、稳定性、决定论的动态）	建立在生物学理论的基础之上（结构、模式、自组织、生命周期）
时间。完全不考虑时间（德布鲁）	时间成为核心
对时间做表面化的处理（增长理论）	时间与年龄紧密联系
基本不考虑年龄	个体生命成为核心因素。年龄就是时间
强调数量、价格与均衡	强调（位置、技术、制度及机会的）结构、模式和功能
经济的元素是数量和价格	经济的元素是模式和机会。在每个社会，相容的结构完成特定功能（参阅人类学）
所用的语言：19世纪的数量和博弈论，不动点拓扑	更定性化的语言。承认博弈论有更多定性用途。其他有用的定性数学工具
实际上没有代际之分	世代更替成为核心。经济体的成员在不断变化、人口的年龄结构也在不断变化。经验“代代相传”

续前表

旧经济学	新经济学
大量使用各种统计指标	关注个体生命。每个人都是独立的、不同的。不断在个体层面与总体层面之间切换。将出现不同的福利指标且只是粗略的度量。个体寿命成为一个指标
因为一切事物都处于均衡，所以不存在真正意义上的动力学。（这类似于周期运动的单摆。）没有真正意义上的变化，而只有“动态的悬浮”	经济永远处于“时间边缘”上，不断向前迈进。结构永远在组合、退化和发展变化。这一切之所以发生，是由于外部性、收益递增、交易成本、结构性排斥所导致的不稳定的运动
如果不存在外部性，同时所有人的能力都相等，那么我们就会到达天堂	外部性和人与人之间的差异成了驱动力量。天堂并不存在。经济系统永远在展开自身
绝大多数问题都无法回答。统一的系统不可兼容	许多问题仍然很难回答。但是假设已经很清晰
“可检验的假说”（萨缪尔森语）。假设存在各种定律	模型拟合数据（例如实验数据）。拟合就是拟合。真正的定律不存在。定律会改变
将研究对象视为结构简单的事物	将研究对象视为天生复杂的事物
把经济学视为“软”的物理学	将经济学视为极其复杂的科学
交换和资源驱动经济发展	外部性、差异、秩序规则、相容性、心智结构、家庭、可能的寿命以及收益递增驱动了制度、社会和经济

引言　复杂性思维造就复杂经济学

[1]　这本文集指的是《经济中的收益递增与路径依赖》（*Increasing Returns and Path Dependence in the Economy*），布莱恩·阿瑟著，美国密歇根大学出版社，1994年出版。

[2]　请参阅约翰·霍兰德和约翰·米勒（John Miller）的文章《经济理论中的人工适应主体》（*Artificial Adaptive Agents in Economic Theory*），载于《美国经济协会会议文章》，1991年。

[3]　请参阅达米安·夏利（Damien Challet）和张翼成（Yi-Cheng Zhang）的文章《协作组织在少数者博弈模型中的出现》，载于《理学A》（*Physica A*），1997年。

[4]　请参阅阿瑟于1993年发表在《科学美国人》第5期上的文章。

[5]　请参阅理查德·伦斯基（Richard Lenski）等人的文章《复杂性特征的进化起源》（*The Evolutionary Origin of Complex Features*），载于《自然》，2003年。

[6]　请参阅阿瑟于1999年发表在《科学》杂志上的文章。

第01章　什么是复杂经济学

[1]　请参阅《作为一个进化的复杂系统的经济》（*The Economy as an Evolving Complex System*）各卷。第一卷的主编是阿罗、安德森和派因斯，于1988年出版；第二卷的主编是阿瑟、布卢默和莱恩，于1997年出版；第三卷的主编是布卢默和杜尔劳夫，于2005年出版。关于“复杂经济学”思想的历史，请参阅丰塔纳（2010）和阿瑟（2010 b）的文章，更通俗一些的解释，请参阅沃尔德罗普（1992）和拜因霍克（2006）的著作。复杂经济学已经有了多种形式的变体，包括生成经济学（generative economics）、互动行为主体经济学（interactive agent economics）、基于行为主体的计算经济学（agent-based computational economics），请参阅爱泼斯坦（2006 a）、米勒和佩奇（2007）、特斯法齐和贾德（2006）的文章。

[2]　关于复杂经济学的一般方法，也请参阅阿克斯特尔（2007）、科兰德（2000，2012）、爱泼斯坦（2006）、法默（2012）、贾协（2006）、基尔曼（2011）、罗瑟（1999）、特斯法齐（2006）等人的文章。“复杂经济学”这个术语是阿瑟于1999年首次使用的，见阿瑟（1999）的文章。

[3]　请参阅布劳格（2003）、布朗克（2009，2011）、卡西第（2009）、科兰德等人（2009）、戴维斯（2007）、法默和格林纲科普洛斯（2008）、基尔曼（2010）、科普尔和卢瑟（2010）、克鲁格曼（2009）、米罗斯基（2002），以及辛普森（2002）的论著。

[4]　1909年，在与熊彼特交谈时，瓦尔拉也曾经表达过类似的观点，“从根本上说，生命是被动的，只有在受到自然或社会的影响时，才会调整自己来适应。因此，关于静态过程的理论，事实上就构成了整个理论经济学”（塔布，1999；莱斯曼，2004）。

[5]　正如凯恩斯在1937年所指出的，“对于欧洲战争的前景、铜的价格以及今后20年的利率等关于这些事件，没有任何科学依据可以让我们得出可计算的概率。我们根本不知道”。

[6]　索罗斯在1987年将此称为“反身性”（reflexivity）准则。

[7]　对此，一种标准的反对性观点认为，如果允许行为主体的推理不是演绎式的，那么就意味着接受任意性的存在。靠什么去阻止行为主体，不会在这种随意的信念或行为中进行特别的选择，以达到某些有利的结果呢？这种情况当然是有可能发生的，但是这并不能证明退回到

那种不切实际的“理性”行为模型的合理性。我们强调的观念是，不能为了简化分析而假设某种行为，而必须假设能够使模型更符合实际的行为。

[8] 事实上，将这种情形称为“有限理性”可以说是用词不当。它意味着行为主体并没有利用自己所掌握的全部推理能力，而这在不确定性情形下往往是错误的。

[9] 当然，“当前状况”也包括相关的历史或对于历史的记忆。

[10] 现代计算思想将计算视为持续的、同时的、并发的、分布的，而且也是概率化的。请参阅美国计算机协会2010年的“普遍存在研讨会”中《什么是计算》（*What Is Computation?*）一文，也可以参阅拜因霍克的论著（2011）。

[11] 在前面，我曾经指出经济的未来是不确定的，严格地说就是经济中并不存在完美的算法规则。因此，在现在这个思想实验中，我假设了一个“上帝”，他可以决定每个行为主体在所有情况下的行为反应。

[12] 包括它是否收敛，或始终停留在某个极限的给定领域内。

[13] 例如，在著名的“爱尔法鲁”问题中，请参阅阿瑟1994年的文章，其中就出现了一个这种类型的生态系统。在那个生态系统中，存在一个全局均衡吸引子状态，但是个体的预测却一直在不断变化。打个比喻，这就像虽然单棵树是一直在变化，但是整个森林的形状却可以保持不变。

[14] 需要注意的是，我们总是可以将任何一个算法模型重写为方程式的模型，因为任何图灵机的计算都可以用方程式来表示，因此严格地说，基于计算机的分析和标准的分析一样都是数学分析。请参阅爱泼斯坦2006年的论著。

[15] 关于计算在数学、物理学、生物学和地球科学等学科中构建理论模型时的作用，请参阅罗伯逊（2003）、贝利（2010）和蔡廷（2006）的文章。

[16] 当然，我们也可以将这个过程建模为一个包括了交通堵塞在内的稳态随机过程，并人为地称之为“均衡”过程。一些新古典模型就是这样做的，请参阅安杰勒托斯（Angeletos）和拉（La）（2011）的文章。这类模型似乎可以否定我的观点，因为我说过标准经济学并不处理非均衡问题。但是，更细致的考察表明，这种非均衡行为一直包含在总体均衡的表象之下，特别包含在一些保持静止的总体随机过程之中，因此被认为处于“均衡”之中。这类模型拓展了“处于均衡”之中的新古典主义范式，但是它们的核心是非均衡过程，因此我将它们纳入到这里来讨论。

[17] 请参阅索罗斯1987年的文章，他认为“股票市场就是各种不同的观点接受检验的地方”。

[18] 关于社会网络，相关文献可谓汗牛充栋。例如，读者可参阅阿尔伯特等人（Albert et al.，2002）、艾伦和盖尔（2000）、梅等人（2008）、纽曼等人（2006）的论著。网络可能是相互“维稳”的，如当一家银行为另外一家银行提供保险时，也可能会相互破坏，如当损失级联性波及整个金融体系时。网络拓扑学研究事件如何快速传播，以及网络互联性是否能提高网络的稳定性，请参阅谢弗尔等人（Scheffer et al.，2012）的文章。

[19] 它们的概率与exp（–|传播长度|）成正比，而不是与大的正态偏差指数exp（–（传播长度）^{2}）成正比。

[20] 在标准经济学中，这些特征之所以不会出现，是因为标准经济学认为行为主体只会对给定的均衡价格做出反应，不会因其他行为主体的行为影响价格的波动而做出反应，因此单个行为主体的随机变化是独立的，是可以加到一起的。这就导致了正态分布。

[21] 因此，这种形式的经济学就被称之为“复杂经济学”。

[22] 在经济学中，关于“中观”（meso）这个术语的早期用法，请参阅多普弗（Dopfer，

2007）、埃尔斯纳和海因里希（Elsner & Heinrich，2009）的文章。

[23] 但是一般而言，这种说法并不正确，许多经济状况并不拥有导致均衡吸引子的力量。

[24] 对于经济形成的研究，还有其他复杂性方法，可以参阅希尔达科（Hildago）和豪斯曼（Hausmann）（2009）、莱恩等人（2009）的文章。关于结构变化，请参阅诺思（1981）的文章。

[25] 熊彼特在1912年认为新产品、新技术组合是经济形成，用他的话来说是“经济发展”的主要动力。

[26] 政治经济学家威廉·塔布（William Tabb），在1999年对结构性变化进行了这样的阐述：“技术革命和政治变革为各种各样的经济可能性创造了条件，而后者又为看似稳定的若干可持续的时期创造了条件。在这些时期中，为适应累积的社会结构变化而设计的监管制度，会呈现出一种有序的进步表象。这些制度形式虽然适合某个发展阶段，但是随着新力量的发展和新的生产关系的出现，就变成了拖累，生机勃勃的市场力量会带来各种各样的社会问题。当这些问题严重到了必须解决的地步时，新的规则、新的监管法规和新的制度就会破壳而出。”

[27] 类似的观点也适用于胚胎发展理论、生化过程理论、分子基因理论，以及细胞生物学理论。例如，有丝细胞分裂模型不涉及任何数学方程，却有一系列便于理解的可能是复杂的阶段和步骤。

[28] 读者也许可以试着将这个理论翻译成我们熟悉的术语，如劳动、资本和增长等。这是有可能做到的，但我更倾向于将它作为“想象”或理解经济变化的另一种有效途径，就像MRI成像仪不同于传统的X射线成像仪一样。

[29] 2006年，沃尔夫冈·波拉克和我成功地利用计算机，构建了一个关于创造过程的模型，在这个模型中，通过对以前的技术，即电路进行组合，越来越复杂的技术，即数字逻辑电路从最初的简单技术中涌现出来。

[30] 关于经济学的这两大“分支”，请参阅塔布的出色的分析（塔布，1999），也请参阅布朗克（2009）的文章。

[31] 请参阅霍尔特等人（2010）和戴维斯（2007）的文章。

第02章　“爱尔法鲁酒吧”问题

[1] 关于这种解决问题的方法的心理学文献，请参阅R.尚克（Schank）和R. P.阿贝尔森（Abelson）（1977）、戴维·鲁梅尔哈特（David Rumelhart）（1980）、戈登·H.鲍尔（Gordon H. Bower）和欧内斯特·希尔加德（Ernest R.Hilgard）（1981），还有约翰·霍兰德等人（1986）的文章。当然，并不是所有决策问题都是以这种方式解决的。绝大多数日常动作，如步行或驾驶，都是在潜意识引导下完成的，即“模式-认知”直接映射为动作。对于这种情况，互联主义模型可以解释。

[2] 类似的结论，也适用于进化博弈论中的各种策略，不过两者之间还是有所不同。在进化博弈论中，并不存在大量个人的主观预期模型，而只有少数策略在相互竞争。

[3] 这让我们想起传奇棒球运动员约吉·贝拉（Yogi Berra）说过的很著名的一句话。当有人问他，为什么不再去圣路易斯著名的鲁杰里（Ruggeri）餐馆吃饭了？他的回答是：“不会再有人去那里了，那里实在太拥挤了。”

第03章　圣塔菲人工股票市场

[1] 这篇文章的一个早期版本，请参阅帕尔默等人（1994）的文章。

[2] 关于这个问题的经典论述，请参阅卢卡斯（1978）的文章，或者迪巴（Diba）和格罗斯曼（Grossman）（1988）的文章。

[3] 有关证据请参阅弗兰克尔（Frankel）和弗鲁特（Froot）（1990）的文章。

[4] 在此不妨引述最成功的交易者之一乔治·索罗斯在1994年所说的一段话："这种有效市场理论对金融市场运行方式的解释，严重歪曲了现实……这个明显不真实的理论，竟然被如此广泛地接受，这实在是一件令人啧啧称奇的事情。"

[5] 有关圣塔菲研究所的人工股票市场的早期报告，见帕尔默等人（1994）的文章。

[6] 还有第二个解，那是一个常数指数增长的解，通常可以通过适当的横截性条件将它排除掉。

[7] 这个术语是汤姆·萨金特在1993年首创的。像我一样，萨金特也认为，在宏观经济背景下，为了形成预期，行为主体需要先成为某种"市场统计学家"。

[8] 将归纳推理应用于资产定价和决策问题的早期研究，请参阅我以前写的几篇文章（阿瑟，1992，1994，1995）。例如，"爱尔法鲁问题"，也请参阅萨金特（1993）的文章。关于心理学以及适应问题的相关文献中对归纳推理的研究，请参阅霍兰德等人（1986）、鲁梅尔哈特（1977），以及尚克和阿贝尔森（1977）的文章。

[9] 在一定意义上，他们是在利用市场数据来学会运用适当的预期模型，并在预期模型之间切换。完美的归纳理性当然也是不确定的。会学习的行为主体的智力可高可低，在不知道他人的学习方法时，无法先验地确定他们的学习方法是不是最有效的。他们要想知道他们的方法的有效性，只有用市场数据来检验。

[10] 在写作本章时，我们发现，如果我们只改变遗传算法的调用速率，那么这两种体制都会出现，而且结果在本质上是相同的。

[11] 在一个更简单的模型中，布卢姆（Blume）和伊斯利（Easley）在1982年从数学上证明了，理性预期均衡的进化稳定性。

[12] 我们也尝试过，当这些实验运行到100万期时，也没有观察到技术交易位消失的迹象。

[13] 自相关的波动率，通常与广义自回归条件异方差时间序列拟合，因此这种行为被称为"广义自回归条件异方差行为"。请参阅波勒斯莱弗等人（1990），古德哈特和奥哈拉（1995）的文章。

[14] 对于不同模型中的波动率聚类的讨论，请参阅尤瑟弗米尔（Youssefmir）和胡伯曼（Huberman）（1995），以及格兰南和斯温德尔（Grannan & Swindle，1994）的文章。

[15] 模型设计时的不同选择会导致不同的结果。我们在评估预期信念的准确性时，所依据的是它们的准确性，而不是它们带来利润的能力。在现实世界的市场中，这种替代方案可能会产生不同的结果。例如，在存在"滑点"的情况下，即指市场一方的交易者很难找到市场另一方的交易者时，如果根据上涨预期买进时用利润而不是用准确性来验证，那么就会得到不同的结果。尽管没有证明，但是我们还是认为，这两个标准会导致相同的定性结果。

[16] 索罗斯在1994年也提出了这样的观点，我们在这里采用了他的"反身性"这个术语。

第04章　收益递增和路径依赖

[1] 见H.金（H. Kim）、A.莫尔斯（A. Morse）和L.辛加勒斯（L. Zing-ales），《自20世纪70年代以来经济学中的重要工作》，载于《经济展望杂志》，2006年，第20卷，第4期。

[2]　罗森伯格（Rosenberg，1982）称之为“干中学”，请参阅阿特金森和斯蒂格利茨（1969）的文章。例如，像波音727这样的喷气式飞机的设计，不仅经历了反复的修改，同时还通过航空公司的飞行，积累了大量实际的使用经验，所以它们在结构健全性、机翼设计、有效载荷能力和发动机效率等方面都有了显著的改进。

[3]　这里使用了一些我引入的术语（阿瑟，1983）。“资助开发的技术”（sponsored technologies）是专有的、能够被定价和策略性地加以操纵的；“无资助技术”则是通用的，不能被操纵或定价的。

[4]　在技术正处于不断改进的时候，使用者在某些情况下“暂且等待一段时间”可能是合算的，但这会导致技术不被采用。请参阅巴尔塞和李普曼（Balcer & Lippman，1984），马默（Mamer）和迈克卡德尔（McCardle）（1987）的文章。为了避免这个问题，我们可以假设采用者需要先采用某种过时的技术，该技术在时间{t_i}分崩离析。

[5]　更现实的情况是，技术带来的货币收益具有不确定性。对于这种情况，我们可以假设一个冯·诺伊曼-摩根斯坦型行为主体，那样一来，表4-1就可以解释为由此而得到的确定的预期效用型收益。

[6]　另一个可以用的效率标准是在n次选择之后的总收益。但是，在目前这个问题中，我们的行为主体有两种类型，且具有不同偏好。根据所谓的“贪婪算法”，每个行为主体都会选择当前对自己来说最好的选择。很容易就可以证明，在任何收益体制下，总收益的最大化从来都是无法保证的。

[7]　在讨论收益递增情况下的商品竞争动力学的文献中，也发现了类似的结果，请参阅斯彭斯（Spence，1981）、弗登伯格（Fudenberg）和梯若尔（Tirole）（1983）的文章。

[8]　关于这项研究，更通俗的一个版本请参阅阿瑟、埃尔莫利耶夫和卡尼奥弗斯基（1987a）的文章。

[9]　这方面的例子还包括，英国铁路采用的窄轨距，请参阅金德尔伯格（Kindleberger，1983）的文章；美国的彩色电视制式；20世纪50年代的编程语言FORTRAN；当然还有QWERTY键盘，请参阅阿瑟（1984）、戴维（1985）、哈特威克（Hartwick，1985）的文章。在这些例子中，收益递增的来源是网络外部性，而不是学习效应。法勒尔（Farrell）和沙隆纳（Saloner）还研究了被锁定的技术标准的突破（1985，1986）。

[10]　有趣的是，迟至1904年，弗莱彻（Fletcher）还说：“……除非汽油动力汽车这些令人反感的特点全都突然消失了，否则它必定会被它那个不令人讨厌的对手，即蒸汽动力汽车赶出市场。”

[11]　对非凸性和路径依赖的重要性的早期研究，请参阅戴维（1975）的文章。

[12]　受资助的技术之间的竞争，受这个缺失的市场的影响较小。提供资助的厂商能够更容易地独占日后的收益，因此他们有动机去开发最初成本昂贵但有前途的技术。此外，为资助技术开发的投资者提供融资的金融市场，以及为那些可能做出“错误”选择的采用者提供保险的保险市场，也能够减轻风险厌恶者的损失。当然，如果一种产品获得了成功并锁定在了市场上，那么也可能会出现垄断定价问题。关于相关政策的进一步讨论，读者可以参阅戴维（1987）的文章。

[13]　对于影响技术的其他选择机制，请参阅多斯（Dosi，1988）、多斯等人（1988），以及梅特卡夫（Metcalfe，1985）的文章。

第05章　经济中的过程与涌现

[1]　因为一个先验的跨期均衡，几乎肯定不能算是一个机制。

[2] 诺曼·帕卡德（Norman Packard）在圣塔菲研究所1987年会议上的报告，谈到了这个“动态系统”方法的问题。他指出：“如果一组相关变量随时间变化，那么状态空间本身就是随时间变化的，这与常规动力系统模型不相称。”

[3] 约翰·霍兰德提交给圣塔菲研究所1987年的会议文章完美地、准确地勾勒了这些特征。关于“圣塔菲方法”的早期描述，读者也可以阅读圣塔菲研究所经济研究计划1989年3月的那篇通讯稿《涌现特征》。

[4] 不过委托代理理论和交易成本经济学除外，在这两种理论中，会出现一个简单的层级结构。

[5] 在这里，我们不需要细究构成经济“单位”和“层次”的到底是什么，因为这因问题情境而异。

[6] 关于交互网络的社会学文献汗牛充栋。入门型文献，读者可以阅读诺里亚（Noria）和埃克勒斯（Eccles）的论著。格兰诺维特（Granovetter）撰写的一篇题为“经济社会学的解释问题”非常值得一读。关于方法论方面的讨论，请参阅沃瑟曼（Wasserman）和福斯特（Faust）的综述。

第06章　再好的经济和社会系统也会被“玩弄”

[1] 关于经济学对2008年金融危机，以及其他危机的“失语”，有关的批判性论著，请参阅科兰德等人（2008）、科普尔和卢瑟（2010）的文章。

[2] 1991年版的《微软词典》，是在第一层含义上使用“剥削”一词的：对现有机会的利用和提炼。该辞典还对“剥削”与“探索”进行了对比，“探索”指不断寻找新的机会。在本文中，“剥削”一词同时包含了这两个要素，因此我们将要讨论的是，行为主体如何探索机会并利用它。

[3] 维基百科在2010年10月9日将“玩弄”定义为：“为了得到自己想要的结果，而‘使用’本来旨在保护系统的规则和程序去操纵系统。”

[4] 请参阅怀特（White，2009）的论著。

[5] 请参阅尼克尔斯（Nichols）和柏林纳（Berliner）（2007）的论著。

[6] 请参阅克里斯特尔（Chrystal）和迈曾（Mizen）（2001）的论著。

[7] 关于我在这里描述的这些步骤的具体实现方法，请参阅阿瑟等人（1997）的文章。

[8] 一种相当符合现实的做法是，我们可以允许行为主体共享哪些策略有效的信息，并允许这种信息在种群中传播。同时，我们可以假设，如果某种策略是有效的，那么人们将会注意到它并生成该策略的各种变体，即他们将会围绕这个策略进行“探索”。

第07章　技术究竟是如何进化的

[1] 奥格本（Ogburn）早在1922年就提出了新的技术由原先存在的组件，即技术构建而成的观点。另外，肯普弗特（Kaempffert）在1930年也指出，新技术是“机械元素的复合体，而这些元素是作为社会遗产的一部分而积累起来的”。关于这个思想的更全面和更严格的论述，请参阅阿瑟的相关文章。

[2] 我们可以说，在“集体意义”上，技术是自我生产的或自创生（auto-poietic）的。“自创生”这个术语是马图拉纳（Maturana）和万利拉（Varela）首创的。当然，技术自身创造了自身这个论断是需要证明的。从根本上说，所有技术都是从对现象的驾驭和利用中创造出来的。但是现象又是通过现有的物理设备和手段，即通过现有技术而被驾驭的。因此，如果我们认同

现象被现有技术利用了，同时我们把创造新技术归入人类活动的类别，那么我们就可以说技术自身创造了自身。

[3] 对于这个网络，更加恰当的定义方法是看什么带来了什么，什么使得什么成为可能，而不能仅仅简单地用每种新技术中包含什么组件来定义。

第08章 技术进化所引发的经济进化

[1] 这个定义取自Dictionary.com网站，WordNet 3.0，美国普林斯顿大学，2008年。

[2] 关于工业革命的历史，请参阅兰德斯（1969）、莫克尔（1990）和阿什顿（1968）等人的文章。

[3] 这句话引用自M. E.罗斯（1981）。也请参阅P. W.巴特里普（Bartrip）和S. B.布曼（Burman），于1983年发表的文章。

[4] 拉普（Rapp）在1989年也提出了一个类似的想法。

第09章 复杂性的进化

[1] 这个引文要感谢丹·麦克谢（Dan McShea）的指点。

[2] 通用电气飞机发动机公司的迈克尔·贝利（Michael Bailey）的个人通信。

[3] 如果它们可以很容易地相互替换的话，载体确实会改变。例如，假设对任何一种基础资产都可以创设期权，那么在这种情况下，载体可以而且确实在快速改变。但是，基础资产的本质属性，即作为一个承载了不确定价值的对象，却在所有情况下都依然是必要的，而且不会改变。

第10章 认知科学

[1] 正文引用的诗句来自爱尔兰诗人叶芝的诗《经柳园而下》（*Down by the Salley Gardens*），是其中的第1、2、7、8句。全诗原文如下：Down by the salley gardens my love and I did meet；She passed the salley gardens with little snow-white feet. She bid me take love easy，as the leaves grow on the tree；But I，being young and foolish，with her would not agree. In a field by the river my love and Idid stand，And on my leaning shoulder she laid her snow-white hand. She bid me take life easy，as the grass grows on the weirs；But I was young and foolish，and now am full of tears.傅浩先生的译文为：我的爱人和我确曾相会在柳园下边，她那一双雪白的小脚款款走过柳园。她让我从容看待爱情，如树头生绿叶，可我，年少无知，不愿听从她的劝诫。我的爱人和我确曾伫立在河畔田间，她那只雪白的小手搭着我斜倚的肩。她让我从容看待人生，如堰上长青草，可我，那时年少无知，如今悔泪滔滔。——译者注

第11章 确定性的终结

[1] 这个例子是在普雷斯科特（Prescott）和维斯克（Visscher）的文章“具有远见的企业的序贯选址决策”的基础上改写的，原文载于《贝尔经济学杂志》，1977年。

主流经济学“挨批评”的历史已经可以写很多本书了，但是总体来看，批评的效果并不算太好。这可能与批评者未能提出一个可以用来替代“旧经济学”框架的“新经济学”框架有关。

布莱恩·阿瑟本人和他的这本书，似乎就是为了完成这个使命而生的。阿瑟是研究收益递增现象的先驱，也是研究技术的本质及技术与经济的关系的先驱，更是引导经济学思维方式的复杂性转向的一个关键的节点人物。阿瑟把他倡导的新经济思想框架称为“复杂经济学”，并把标准经济学作为它的一个特例。复杂经济学认为：经济未必总处于均衡状态；在同一种经济状况下，既可能出现收益递增，也可能出现收益递减；经济系统本身不是给定的，而是在一系列制度、规范和技术创新的不断进化中形成的；个体的行动和战略都处于不断的进化当中，结构随着时间流逝不断形成且不断重组；世界是一个有机的、进化的、充斥着历史偶然性的世界；创新，即经济、技术和制度的变迁不一定是连续的，而是断裂的，新的世界也许会突然取代旧的世界……

复杂经济学主要起源于20世纪80年代的圣塔菲研究所，它的第一个“复杂性”研究项目，就是阿瑟主持的。这本书收录的这些文章，时间跨度超过了20年。从这些文章中，不仅可以看出阿瑟本人的思想发展轨迹，而且可以了解经济学思想史（科学史）上的一个重要发展时期的一个侧面。

现在，“复杂性”或“复杂系统”一词已经有些被用滥了，阿瑟此书也可起到正本清源的作用。确实，任何一个人，如果真的对复杂性现象、对作为复杂系统的经济感兴趣，阿瑟都是一个他绕不开的人物。

感谢我的妻子傅瑞蓉，她是本译稿的第一位读者和批评者，帮助我改正了不少错误。感谢儿子贾岚晴带给我的快乐。感谢岳父傅美峰、岳母蒋仁娟对贾岚晴的悉心照料。借此机会，还要感谢汪丁丁教授、叶航教授和罗卫东教授的教诲。感谢何永勤、虞伟华、余仲望、鲍玮玮、傅晓燕、傅锐飞、傅旭飞、陈叶烽、李欢、丁玫、何志星、陈贞芳、楼霞、郑文英、商瑜、李晓玲等好友的帮助。

感谢简学，他一再热情邀请我翻译此书，希望我的译文不致令他失望。书中错漏之处在所难免，敬请专家和读者批评指正。

贾拥民

第01章　什么是复杂经济学

ACM Ubiquity Symposium: *What Is Computation?* 2010.

Angeletos, G-M., and J. La'O, “Decentralization, Communication, and the Origins of Fluctuations,” NBER Working Paper 17060, NBER, Cambridge, MA, 2011.

Albert, R., H. Jeong, and A-L. Barabasi, “Attack and Error Tolerance of Complex Networks,” *Nature*, 406, 379–382, 2000.

Allen, F., and D. Gale, “Financial Contagion,” *Journal of Political Economy*, 108, 1, 1–33, 2000.

Arrow, K., P. Anderson, and D. Pines, *The Economy as an Evolving Complex System*, Addison-Wesley, Reading, MA, 1988.

Arthur, W. B., “Competing Technologies, Increasing Returns, and Lock-In by Historical Events,” *Economic Journal*, 99, 116–131, 1989.

Arthur, W. B., “Bounded Rationality and Inductive Behavior (the El Farol problem),” *American Economic Review Papers and Proceedings*, 84, 406–411, 1994a.

Arthur, W. B., *Increasing Returns and Path Dependence in the Economy*, University of Michigan Press, Ann Arbor, MI, 1994b.

Arthur, W. B., S. Durlauf, and D. Lane, eds. *The Economy as an Evolving Complex System II*, Addison-Wesley, Reading, MA, 1997.

Arthur, W. B., J. H. Holland, B. LeBaron, R. Palmer, and P. Tayler, “Asset Pricing under Endogenous Expectations in an Artificial Stock Market,” in *The Economy as an Evolving Complex System II*, (*op. cit.*), 1997.

Arthur, W. B., “Complexity and the Economy,” *Science*, 284, 107–109, 1999. Arthur, W. B., “Out-of-equilibrium Economics and Agent-based Modeling,” in L. Tesfatsion and K. Judd, (*op. cit.* below), 2006.

Arthur, W. B., and W. Polak, “The Evolution of Technology in a

Simple Computer Model," *Complexity*, 11, 5, 2006.

Arthur, W. B., *The Nature of Technology*: *What It Is and How It Evolves*, The Free Press, New York, 2009.

Arthur, W. B., "Exploitive Behavior in Policy Systems," Mss., IBM Almaden, 2010a.

Arthur, W. B., "Complexity, the Santa Fe Approach, and Nonequilibrium Economics," *Hist. Econ. Ideas*, 18, 2, 149–166, 2010b.

Axtell, R., "What Economic Agents Do: How Cognition and Interaction Lead to Emergence and Complexity," *Rev. Austrian Econ.*, 20, 105–122, 2007.

Bailey, J., *Emerge: The Data-Rich Mathematical Infinitesimals of Life, MAPematics*, The Mapematics Institute, 2010.

Beinhocker, E., *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*, Harvard Business School Press, Cambridge, MA, 2006.

Beinhocker, E., "Evolution as Computation: Integrating Self-Organization with Generalized Darwinism," *J. Inst. Econ.*, 7, 3, 393–423, 2011.

Blaug, M., "The Formalist Revolution of the 1950s," *J. History of Economic Thought*, 25, 2, 145–156, 2003.

Blume, L., and S. Durlauf, *The Economy as an Evolving Complex System III*, Oxford University Press, New York, 2006.

Brock, W. A., J. Lakonishok, and B. LeBaron, "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," *Journal of Finance*, 47, 1731–1764, 1992.

Bronk, R., *The Romantic Economist: Imagination in Economics*, Cambridge University Press, Cambridge, UK, 2009.

Bronk, R., "Epistemological Difficulties with Neoclassical Economics," *Rev. Austrian Economics*, 2014.

Cassidy, J., *How Markets Fail: The Logic of Economic Calamities*, Farrar, Straus and Giroux, New York, 2009.

Chaitin, G., *Meta Math! The Quest for Omega*, Vintage Books, New York, 2006.

Chaitin, G., *Proving Darwin: Making Biology Mathematical*, Pantheon Books, New York, 2012.

Colander, D., and R. Kupers. *Laissez-Faire Activism: The Complexity Frame for Policy*, Princeton University Press, Princeton, NJ, 2014.

Colander, D. (ed.), *The Complexity Vision and the Teaching of Economics*, E. Elgar, Cheltenham, UK, 2000.

Colander, D., M. Goldberg, A. Haas, K. Juselius, T. Lux, H. Föllmer, A. Kirman, and B. Sloth, "The Financial Crisis and the Systemic Failure of the Economics Profession," *Critical Rev.*, 21, 2, 2009.

Davis, J., "The Turn in Recent Economics and Return of Orthodoxy," *Cambridge J. Econ.,* 32, 349–366, 2008.

Dopfer, K., "The Pillars of Schumpeter's Economics: Micro, Meso, Macro," in Hanusch and Pyka, eds., *Elgar Companion to Neo-Schumpeterian Economics*. E. Elgar, Cheltenham, UK, 2007.

Elsner, W., and T. Heinrich, "A Simple Theory of 'Meso: ' On the Co-evolution of Institutions and Platform Size," in *Journal of Socio-Economics*, 38, 843–858, 2009.

Epstein J., *Generative Social Science*, Princeton University Press, Princeton, NJ, 2006a.

Epstein, J., "Remarks on the Foundations of Agent-based Generative Social Science," in Tesfatsion and Judd, eds., *op. cit.* (below), 2006b.

Farmer, J. D., and J. Geanakoplos, "The Virtues and Vices of Equilibrium and the Future of Financial Economics," *Complexity*, 14, 8, 11–38, 2008.

Farmer, J. D., "Economics Needs to Treat the Economy as a Complex System," Mss., 2012.

Fontana, M., "Can Neoclassical Economics Handle Complexity?

The Fallacy of the Oil Spot Dynamic," *J. Economic Behavior & Organization*, 76, 584–596, 2010.

Fontana, M. "The Santa Fe Perspective on Economics," *Hist. Econ. Ideas*, 18, 2, 167–196, 2010.

Galla, T., and J. D. Farmer, "Complex Dynamics in Learning Complicated Games," Mss., 2012.

Haldane, A. G., "Rethinking the Financial Network," Speech at Financial Student Assoc., Amsterdam. Bank of England, 2009.

Harris, D. J., "Joan Robinson on 'History versus Equilibrium, ' " J. Robinson Centennial Conference, Burlington, VT, 2003.

Hildago, C. A., and R. Hausmann, "The Building Blocks of Economic Complexity," *Proc. Nat. Acad. Sci.*, 106, 26, 10570–10575, 2009.

Holland, J., K. Holyoak, R. Nisbett, and P. Thagard, *Induction*, MIT Press, Cambridge, MA, 1986.

Holt R., J. Rosser, and D. Colander, "The Complexity Era in Economics," Middlebury College, 2010.

Hommes, C. H. "Bounded Rationality and Learning in Complex Markets," in Rosser, J. B., (*op. cit.* below), 2009.

Judd, K., "Computationally Intensive Analyses in Economics," in Tesfatsion and Judd, eds., (*op. cit. below*), 2006.

Keynes, J. M., "The General Theory of Employment," *Quarterly Journal of Economics*, 51, 209–233, 1937.

Kirman, A., *Complex Economics*, Routledge, New York, 2011.

Kirman, A., "The Economic Crisis Is a Crisis for Economic Theory," *CESifo Economic Studies*, 56, 4, 498–535, 2010.

Kopel, M., "Oligopoly Dynamics," in Rosser, J. B., (*op. cit.* below), 2009.

Koppl, R., and W. Luther, "BRACE for a New Interventionist Economics," Fairleigh Dickinson University, 2010.

Krugman, P., “How Did Economists Get It So Wrong?” *New York Times*, Sept. 6,2009.

Lane, D., D. Pumain, S. van der Leeuw, and G. West (eds.), *Complexity Perspectives in Innovation and Social Change*, Springer, Berlin, 2009.

LeBaron, B., W. B. Arthur, and R. Palmer, “Time Series Properties of an Artificial Stock Market, *Journal of Econ. Dynamics and Control*, 23, 1487–1516, 1999.

Lindgren, K., “Evolutionary Phenomena in Simple Dynamics,” in C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, eds, *Artificial Life II*, Addison-Wesley, Reading, MA, 1991.

Louca, F., “Bounded Heresies: Early Intuitions of Complexity in Economics.” *Hist. Econ. Ideas*, 18, 2, 77–113, 2010.

May, R., S. Levin, G. Sugihara, “Complex Systems: Ecology for Bankers,” *Nature*, 451, 893–895, 2008.

Miller, J., and S. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, Princeton University Press, Princeton, NJ, 2007.

Mirowski, P., *Machine Dreams: Economics Becomes a Cyborg Science*, Cambridge University Press, Cambridge, UK, 2002.

Newman, M., A-L. Barabasi, and D. Watts (eds.). *The Structure and Dynamics of Networks*, Princeton University Press, Princeton, NJ, 2006.

North, D., *Structure and Change in Economic Theory*, Norton, New York, 1981.

Palmer, R. G., W. B. Arthur, J. Holland, B. LeBaron, P. Tayler, “Artificial Economic Life: A Simple Model of a Stock Market,” *Physica D*, 75, 264–274, 1994.

Perez, C., *Technological Revolutions and Financial Capital*, E. Elgar, Cheltenham, UK, 2002.

Reisman, D., *Schumpeter's Market: Enterprise and Evolution*, E. Elgar, Cheltenham, UK, 2004.

Robinson, J., "Time in Economic Theory," *Kyklos*, 33, 2, 219–229, 1980.

Robinson, J., "What Has Become of the Keynesian Revolution?" in J. Robinson,ed. *After Keynes*, Basil Blackwell, Oxford, 1973.

Robertson, D. S. *Phase Change: The Computer Revolution in Science and Mathematics*, Oxford, New York, 2003.

Rosser, J. B., "On the Complexities of Complex Economic Dynamics," *J. Econ. Perspectives*, 13, 4, 1999.

Rosser, J. B., ed. *Handbook of Research on Complexity*, E. Elgar, Cheltenham, UK, 2009.

Samuelson, P. A., *Foundations of Economic Analysis*, Harvard University Press, Cambridge, MA, 1983 (originally 1947).

Sargent T. J., *Bounded Rationality in Macroeconomics*, Clarendon Press, Oxford, UK, 1993.

Scheffer, M., S. Carpenter, T. Lenton, J. Bascompte, W. Brock, V. Dakos, J. van de Koppel, I. van de Leemput, S. Levin, E. van Nes, M. Pascual, and J. Vandermeer,"Anticipating Critical Transitions," *Science*, 338, 19 Oct. 2012.

Schumpeter, J. A., *Das Wesen und der Hauptinhalt der theoretischen Nationalö-konomie*. Dunker & Humbolt, Leipzig, 1908.

Schumpeter, J. A., *The Theory of Economic Development* (1912), Oxford University Press, London, 1961.

Schumpeter, J. A., *History of Economic Analysis*, Allen and Unwin, London, 1954.

Shackle, G. L. S., *Uncertainty in Economics*, Cambridge University Press, Cambridge, UK, 1955.

Shackle, G. L. S., *Epistemics and Economics*, Transaction Publishers, Piscataway, NJ, 1992.

Simpson, D., *Rethinking Economic Behaviour*, St. Martin's Press, New York, 2002.

Smolin, L., "Time and Symmetry in Models of Economic Markets,"

Mss., Feb 2009.

Smolin, L., *Time Reborn*. Houghton, Mifflin, Harcourt, New York, 2013.

Solow, R., “Technical Change and the Aggregate Production Function,” *Rev. Economics and Statistics*, 39, 312–320, 1957.

Soros, G., *The Alchemy of Finance*, Simon & Schuster, New York, 1987.

Tabb, W., *Reconstructing Political Economy*, Routledge, New York, 1999.

Tesfatsion, L., and K. L. Judd, eds., *Handbook of Computational Economics: Vol.2. Agent-Based Computational Economics*, North-Holland Elsevier, New York, 2006.

Tesfatsion, L., “Agent-Based Computational Economics: A Constructive Approach to Economic Theory,” in Tesfatsion and Judd, *op. cit.*, 2006.

Turing, A. M., “On Computable Numbers, with an Application to the Entscheidungsproblem,” *Proc. London Math. Society*, Series 2, 42, 1936.

Waldrop, M., *Complexity*, Simon & Schuster, New York, 1992.

Watts, D., “A Simple Model of Global Cascades on Random Networks,” *PNAS*, 9, 5766–5771, 2002.

Wolfram, S., *A New Kind of Science*, Wolfram Media, Champaign, IL, 2002.

第02章 “爱尔法鲁酒吧”问题

Arthur, W. Brian. “On Learning and Adaptation in the Economy.” Santa Fe Institute (Santa Fe, NM) Paper 92-07-038, 1992.

Bower, Gordon H., and Hilgard, Ernest R. *Theories of learning*. Englewood Cliffs, NJ: Prentice-Hall, 1981.

De Groot, Adriann. *Thought and choice in chess*, Psychological Studies, No. 4. Paris: Mouton, 1965.

Feldman, Julian. “Computer Simulation of Cognitive Processes,” in Harold Borko, ed., *Computer applications in the behavioral sciences*. Englewood Cliffs, NJ: Prentice-Hall, 1962, pp. 336-359.

Grannan, E. R., and Swindle, G. H. “Contrarians and Volatility Clustering.” Santa Fe Institute (Santa Fe, NM) Paper 94-03-010, 1994.

Holland, John H., Holyoak, Keith J., Nisbett, Richard E., and Thagard, Paul R. *Induction*. Cambridge, MA: MIT Press, 1986.

Koza, John. *Genetic programming*. Cambridge, MA: MIT Press, 1992.

Rumelhart, David. “Schemata: The Building Blocks of Cognition,” in R. Spiro,B. Bruce, and W. Brewer, eds., *Theoretical issues in reading comprehension*. Hillside, NJ: Erlbaum, 1980, pp. 33–58.

Sargent, Thomas J. *Bounded rationality in macroeconomics*. New York: Oxford University Press, 1994.

Schank, R., and Abelson, R. P. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillside, NJ: Erlbaum, 1977.

第03章　圣塔菲人工股票市场

W. B. Arthur (1992), “On Learning and Adaptation in the Economy,” Working Paper 92-07-038, Santa Fe Institute, Santa Fe, NM.

W. B. Arthur (1994), “Inductive Behavior and Bounded Rationality,” *Amer. Econ.Rev.*, 84, pp. 406–411.

W. B. Arthur (1995), “Complexity in Economic and Financial Markets,” *Complexity*, 1, pp. 20–25.

A. Beltratti-S. Margarita (1992), “Simulating an Artificial Adaptive Stock Market,” mimeo, Turin University.

L. Blume-D. Easley (1982), “Learning to Be Rational,” *J. Econ. Theory*, 26, pp.340–351.

L. Blume-D. Easley (1990), “Evolution and Market Behavior,” *J.*

Econ. Theory, 58, pp. 9–40.

T. Bollerslev-R. Y. Chou-N. Jayaraman-K. F. Kroner (1990), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *J. Econometrics*, 52, pp. 5–60.

P. Bossaerts (1995), "The Econometrics of Learning in Financial Markets," *Econometric Theory*, 11, pp. 151–189.

M. Bray (1982), "Learning, Estimation, and Stability of Rational Expectations," *J. Econ. Theory*, 26, pp. 318–339.

W. Brock-J. Lakonishok-B. LeBaron (1991), "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," Working Paper 91-01-006, Santa Fe Institute, Santa Fe, NM.

W. Brock-C. H. Hommes (1996), "Models of Complexity in Economics and Finance," mimeo, Department of Economics, University of Wisconsin, Madison.

D. M. Cutler-J. M. Poterba-L. H. Summers (1989), "What Moves Stock Prices?," *J. Portfolio Mgmt.*, 15, pp. 4–12.

J. B. De Long-A. Shleifer-L. H. Summers-R. J. Waldmann (1990a), "Noise Trader Risk in Financial Markets," *J. Pol. Econ.*, 98, pp. 703–738.

J. B. De Long-A. Shleifer-L. H. Summers-R. J. Waldmann (1990b), "Positive Feedback Strategies and Destabilizing Rational Speculation," *J. Fin.*, 45, pp. 379–395.

J. B. De Long-A. Shleifer-L. H. Summers-R. J. Waldmann (1991), "The Survival of Noise Traders in Financial Markets," *J. Bus.*, 64, pp. 1–18.

B. T. Diba-H. I. Grossman (1988), "The Theory of Rational Bubbles in Stock Prices," *Econ. Jour.*, 98, pp. 746–754.

M. Eigen-P. Schuster (1979), *The Hypercycle: A Principle of Natural Self-Organization*, Springer, New York.

E. F. Fama-K. R. French (1988), "Permanent and Temporary Components of Stock Market Prices," *J. Pol. Econ.*, 96, pp. 246–273.

J. A. Frankel-K. A. Froot (1990), “Chartists, Fundamentalists, and Trading in the Foreign Exchange Market,” *AEA Papers & Proc.*, 80, pp. 181–185.

D. Friedman-M. Aoki (1992), “Inefficient Information Aggregation as a Source of Asset Price Bubbles,” *Bull. Econ. Res.*, 44, pp. 251–279.

C. A. E. Goodhart-M. O’Hara (1995), “High Frequency Data in Financial Markets: Issues and Applications,” Unpublished manuscript, London School of Economics.

E. R. Grannan-G. H. Swindle (1994), “Contrarians and Volatility Clustering,” Working Paper 94-03-010, Santa Fe Institute, Santa Fe, NM.

S. J. Grossman (1976), “On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information,” *J. Fin.*, 31 pp. 573–585.

S. J. Grossman-J. Stiglitz (1980), “On the Impossibility of Informationally Efficient Markets,” *Amer. Econ. Rev.*, 70, pp. 393–408.

J. H. Holland-K. J. Holyoak-R. E. Nisbett-P. R. Thagard (1986), *Induction*, MIT Press, Cambridge, MA.

S. Kauffman (1993), *The Origin of Order: Self-Organization and Selection in Evolution*, Oxford University Press, New York.

J. M. Keynes (1936), *General Theory of Employment, Interest, and Money,* Macmillan, London.

A. Kirman (1991), “Epidemics of Opinion and Speculative Bubbles in Financial Markets,” in *Money and Financial Markets*, M. Taylor (ed.), Macmillan, London.

A. W. Kleidon (1986), “Anomalies in Financial Economics: Blueprint for Change?,” *J. Bus.*, 59, pp. 285–316 (supplement).

M. Kurz (1994), “On the Structure and Diversity of Rational Beliefs,” *Econ. Theory*, 4, pp. 877–900.

M. Kurz (1995), “Asset Prices with Rational Beliefs,” Working Paper 96-003, Economics Department, Stanford University.

S. F. Leroy-R. D. Porter (1981), “Stock Price Volatility: Tests Based on Implied Variance Bounds,” *Econometrica*, 49, pp. 97–113.

A. W. Lo-C. Mackinlay (1988), “Stock Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test,” *Rev. Fin. Stud.*, 1, pp. 41–66.

R. E. Lucas (1978), “Asset Prices in an Exchange Economy,” *Econometrica*, 46, pp. 1429–1445.

L. Marengo-H. Tordjman (1995), “Speculation, Heterogeneity, and Learning: A Model of Exchange Rate Dynamics,” Working Paper WP-95-17, IIASA.

P. Milgrom-N. Stokey (1982), “Information, Trade, and Common Knowledge,” *J. Econ. Theory*, 26, pp. 17–27.

R. Nelson-S. Winter (1982), *An Evolutionary Theory of Economic Change,* Harvard University, Press/Bellknap, Cambridge, MA.

R. G. Palmer-W. B. Arthur-J. H. Holland-B. Lebaron-P. Tayler (1994), “Artificial Economic Life: A Simple Model of a Stockmarket,” *Physica* D, 75, pp. 264–274.

J. M. Poterba-L. H. Summers (1988), “Mean Reversion in Stock Prices: Evidence and Implications,” *J. Fin. Econ.*, 22, pp. 27–59.

C. Rieck (1994), “Evolutionary Simulation of Asset Trading Strategies,” in *Many-Agent Simulation and Artificial Life*, E. Hillenbrand and J. Stender (eds.), IOS Press, Washington, DC.

D. E. Rumelhart (1977), *Human Information Processing*, Wiley, New York.

T. J. Sargent (1993), *Bounded Rationality in Macroeconomics*, Oxford University Press, New York.

R. Schank-R. P. Abelson (1977), *Scripts, Plans, Goals, and Understanding,* Erlbaum, Hillsdale, NJ.

R. J. Shiller (1981), “Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?,” *Amer. Econ. Rev.*, 71, pp. 421–436.

R. J. Shiller (1989), *Market Volatility*, MIT Press, Cambridge, MA.

A. Shleifer-L. H. Summers (1990), "The Noise Trader Approach to Finance," *J. Econ. Perspectives*, 4, pp. 19–33.

G. Soros (1994), *The Theory of Reflexivity*, Soros Fund Management, New York.

L. H. Summers (1986), "Does the Stock Market Rationally Reflect Fundamental Values?," *J. Fin.*, 46, pp. 591–601.

R. H. Thaler (1993), *Advances in Behavioral Finance*, Russell Sage Foundation, New York.

M. Youssefmir-B. Huberman (1995), "Clustered Volatility in Multiagent Dynamics," Working Paper 95-05- 051, Santa Fe Institute, Santa Fe, NM.

第04章 收益递增和路径依赖

Agnew, H. (1981). "Gas-cooled nuclear power reactors." *Scientific American*, vol. 244, pp. 55–63.

Arthur, W. B. (1983). "Competing technologies and lock-in by historical small events: the dynamics of allocation under increasing returns." International Institute for Applied Systems Analysis Paper WP-83-92, Laxenburg, Austria. (Center for Economic Policy Research, Paper 43, Stanford).

Arthur, W. B. (1984). "Competing technologies and economic prediction." *Options*, International Institute for Applied Systems Analysis, Laxenburg, Austria. No.1984/2, pp. 10–13.

Arthur, W. B., Ermoliev, Yu, and Kaniovski, Yu. (1983). "On generalized urn schemes of the polya kind." *Kibernetika*, vol. 19, pp. 49–56. English translation in *Cybernetics*, vol. 19, pp. 61–71.

Arthur, W. B., Ermoliev, Yu, and Kaniovski, Yu. (1986). "Strong laws for a class of path-dependent urn processes." In *Proceedings of the International Conference on Stochastic Optimization, Kiev* 1984, Springer Lecture Notes Control and Information Sciences, vol. 81, pp. 187–300.

Arthur, W. B., Ermoliev, Yu, and Kaniovski, Yu. (1987*a*). "Path-dependent processes and the emergence of macro-structure." *European Journal of Operational Research*, vol. 30, pp. 294–303.

Arthur, W. B., Ermoliev, Yu, and Kaniovski, Yu. (1987*b*). "Non-linear urn processes: asymptotic behavior and applications." International Institute for Applied Systems Analysis. Paper WP-87-85, Laxenburg, Austria.

Arthur, W. B., Ermoliev, Yu, and Kaniovski, Yu. (1988). "Self-reinforcing mechanisms in economics." In *The Economy as an Evolving Complex System* (P. Anderson, K. Arrow, and D. Pines (eds)). Reading, Massachusetts: Addison–Wesley.

Atkinson, A. and Stiglitz, J. (1969). "A new view of technical change." Economic Journal, vol. 79, pp. 573–580.

Balcer, Y., and Lippman, S. (1984). "Technological expectations and the adoption of improved technology," *Journal of Economic Theory*, vol. 34, pp. 292–318.

Bupp, I., and Derian, J. (1978). *Light Water: How The Nuclear Dream Dissolved*. New York: Basic.

Burton, R. (1976). "Recent advances in vehicular steam engine efficiency." Society of Automotive Engineers, Preprint 760340.

Cowan, R. (1987). "Backing the wrong horse: sequential choice among technologies of unknown merit." Ph. D. Dissertation, Stanford.

David, P. (1975). *Technical Choice, Innovation, and Economic Growth*. Cambridge: Cambridge University Press.

David, P. (1985). "Clio and the economics of QWERTY." *American Economic Review Proceedings*, vol. 75, pp. 332–337.

David, P. (1987). "Some new standards for the economics of standardization in the information age." In *Economic Policy and Technological Performance*, (P. Dasgupta and P. Stoneman, eds.) Cambridge: Cambridge University Press.

David, P., and Bunn, J. (1987). "The economics of gateway technologies and network evolution: lessons from electricity supply

history.” Centre for Economic Policy Research, Paper 119, Stanford.

Dosi, G. (1988). “Sources, procedures and microeconomic effects of innovation.” *Journal of Economic Literature*, vol. 26, pp. 1120–1171.

Dosi, G., Freeman, C., Nelson, R., Silverberg, G. and Soete, L. (eds.) (1988). *Technical Change and Economic Theory*, London: Pinter.

Farrell, J., and Saloner, G. (1985). “Standardization, compatibility, and innovation.” *Rand Journal of Economics*, vol. 16, pp. 70–83.

Farrell, J., and Saloner, G. (1986). “Installed base and compatibility: innovation, product preannouncements and predation.” *American Economic Review*, vol. 76. pp.940–955.

Fletcher, W. (1904). *English and American Steam Carriages and Traction Engines* (reprinted 1973). Newton Abbot: David and Charles.

Fudenberg, D., and Tirole, J. (1983). “Learning by doing and market performance.” *Bell Journal of Economics*, vol. 14, pp. 522–530.

Hanson, W. (1985). “Bandwagons and orphans: dynamic pricing of competing systems subject to decreasing costs.” Ph. D. Dissertation, Stanford.

Hartwick, J. (1985). “The persistence of QWERTY and analogous suboptimal standards.” Mimeo, Queen's University, Kingston, Ontario.

Hill, B., Lane, D. and Sudderth, W. (1980). “A strong law for some generalized urn processes.” *Annals of Probability*, vol. 8, pp. 214–226.

Katz, M., and Shapiro, C. (1985). “Network externalities, competition, and compatibility.” *American Economic Review*, vol. 75, pp. 424–440.

Katz, M., and Shapiro, C. (1986). “Technology adoption in the presence of network externalities.” *Journal of Political Economy*, vol. 94, pp. 822–841.

Kindleberger, C. (1983). “Standards as public, collective and private goods.” *Kyklos*, vol. 36, pp. 377–396.

Mamer, J., and McCardle, K. (1987). “Uncertainty, competition

and the adoption of new technology.” *Management Science*, vol. 33, pp. 161–177.

McLaughlin, C. (1954). “The Stanley Steamer: a study in unsuccessful innovation.” *Explorations in Entrepreneurial History*, vol. 7, pp. 37–47.

Metcalfe, J. S. (1985). “On technological competition.” Mimeo, University of Manchester.

Rosenberg, N. (1982). *Inside the Black Box: Technology and Economics*. Cambridge: Cambridge University Press.

Strack, W. (1970). “*Condensers and Boilers for Steam-powered Cars*.” NASA Technical Note, TN D-5813, Washington, D. C.

Spence, A. M. (1981). “The learning curve and competition.” *Bell Journal of Economics*, vol. 12, pp. 49–70.

第05章　经济中的过程与涌现

Baker, W. “The Social Structure of a National Securities Market.” *Amer. J.Sociol.* 89, (1984): 775–811.

Clark, A. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press, 1997.

Eigen, M., and P. Schuster. *The Hypercycle*. Berlin: Springer Verlag, 1979.

Fontana, W., and L. Buss. “The Arrival of the Fittest: Toward a Theory of Biological Organization.” *Bull. Math. Biol.* 56 (1994): 1–64.

Holland, J. H. “The Global Economy as an Adaptive Process.” In *The Economy as an Evolving Complex System*, edited by P. W. Anderson, K. J. Arrow, and D. Pines, 117–124. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. V. Redwood City, CA: Addison-Wesley, 1988.

Hutchins, E. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.

Noria, N., and R. Eccles (Eds.) *Networks and Organizations:*

Structure, Form, and Action. Cambridge, MA: Harvard Business School Press, 1992.

Wasserman, W., and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press, 1994.

第06章　再好的经济和社会系统也会被“玩弄”

Appleton, Michael, “SEC Sues Goldman over Housing Market Deal.” *New York Times*, Apr. 16, 2010.

Arrow, Kenneth, “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review*, 53: 91–96, 1963.

Arthur, W. Brian. “Bounded Rationality and Inductive Behavior (the El Farol problem),” *American Economic Review Papers and Proceedings*, 84, 406–411, 1994.

Arthur, W. Brian, J. H. Holland, B. LeBaron, R. Palmer, and P. Tayler, “Asset Pricing under Endogenous Expectations in an Artificial Stock Market,” in *The Economy as an Evolving Complex System II*, Arthur, W. B., Durlauf, S., Lane, D., eds. Addison-Wesley, Redwood City, CA, 1997.

Axelrod, Robert. *The Evolution of Cooperation*. Basic Books, New York, 1984.

Bibel, George. *Beyond the Black Box: The Forensics of Airplane Crashes*. Johns Hopkins University Press, Baltimore, MD, 2008.

Boyes, Roger. *Meltdown Iceland: How the Global Financial Crisis Bankrupted an Entire Country*. Bloomsbury Publishing, London, 2009.

Campbell, Donald, “Assessing the Impact of Planned Social Change,” Public Affairs Center, Dartmouth, NH, Dec. 1976.

Cassidy, J., *How Markets Fail: the Logic of Economic Calamities*. Farrar, Straus and Giroux, New York, 2009.

Chrystal, K. Alec, and Paul Mizen, “Goodhart's Law: Its Origins, Meaning and Implications for Monetary Policy,” (http://cyberlibris.typepad.com/blog/files/Goodharts_Law.pdf), 2001.

Colander, David, A. Haas, K. Juselius, T. Lux, H. Follmer, M. Goldberg, A. Kirman, B. Sloth, “The Financial Crisis and the Systemic Failure of Academic Economics,” mimeo, 98th Dahlem Workshop, 2008.

Holland, John. *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA, 1992. (Originally published 1975.)

Holland, John H., K. J. Holyoak, R. E. Nisbett and P. R. Thagard, *Induction*. Cambridge, MA, MIT Press, 1986.

Jonsson, Asgeir. *Why Iceland? How One of the World's Smallest Countries Became the Meltdown's Biggest Casualty*, Mc-Graw-Hill, New York, 2009.

Koppl, Roger, and W. J. Luther, “BRACE for a new Interventionist Economics,” mimeo, Fairleigh Dickinson University, 2010.

Lindgren, Kristian. “Evolutionary Phenomena in Simple Dynamics,” in C. Langton, C. Taylor, D. Farmer, S. Rasmussen, (eds.), *Artificial Life II*. Addison-Wesley, Reading, MA, 1992.

Mahar, Maggie. *Money-driven Medicine*. HarperCollins, New York, 2006.

March, James, “Exploration and Exploitation in Organizational Learning,” *Organization Science*, 2, 1, 71–87, 1991.

Morgenson, Gretchen (*New York Times* reporter), “Examining Goldman Sachs,” NPR interview in *Fresh Air*, May 4, 2010.

Nichols, S. L., and D. Berlner, *Collateral Damage*: *How High-stakes Testing Corrupts America's Schools*. Harvard Education Press, Cambridge, MA, 2007.

Pew Charitable Trusts, “History of Fuel Economy: One Decade of Innovation, Two Decades of Inaction” (www. pewtrusts. org), 2010.

Schneier, Bruce, “Cryptography: The Importance of Not Being Different,” *IEEE Computer*, 32, 3, 108–109, 1999.

Suderman, Peter, “Quit Playing Games with my Health Care System,” *Reason*, April 5, 2010.

Tabb, William, *The Restructuring of Capitalism in our Time*. Columbia University Press, New York, 2012.

White, Lawrence J. “The Credit Rating Agencies and the Subprime Debacle.” *Critical Review*, 21 (2–3): 389–399, 2009.

Zuill, Lilla, “AIG's Meltdown Has Roots in Greenberg Era.” *Insurance Journal*, 87, March 3, 2009.

第07章　技术究竟是如何进化的

Schumpeter, J. A. *The Theory of Economic Development*; Transaction Publishers: New Brunswick, NJ, 1911/1934/1996.

Lenski, R. E.; Ofria, C.; Pennock, R. T.; Adami C. The evolutionary origin of complex features. *Nature* 2003, 423, 139–143.

Beck-Sickinger, A.; Weber, P. *Combinatorial Strategies in Biology and Chemistry*; Wiley: Chichester, England, 2001.

Kauffman, S. A. *Investigations*; Oxford University Press: New York, 2002.

Bak, P.; Wiesenfeld, K. Self-organized criticality: An explanation for *1/f noise*. *Phys Rev A* 1988, 38, 364.

Bryant, R. E. Graph-based algorithms for Boolean function manipulation. *IEEE Trans Comput* 1988, 35 (8), 677–691.

Bryant, R. E. Symbolic Boolean manipulation with ordered binary-decision diagrams. *ACM Comput Surv* 1992, 24 (3), 293–318.

Ogburn, W. F. *Social Change*; 1922 (Dell: New York, 1966).

Kaempffert, W. *Invention and Society*. Reading with a Purpose Series, 56, American Library Association: Chicago, 1930.

Arthur, W. B. *The Nature of Technology: What It Is and How It Evolves*; in preparation. (Appeared in 2009, The Free Press: New York.)

Maturana H.; Varela F. Autopoiesis and cognition: The realization of the living. *Boston Studies in the Philosophy of Science* 42; Cohen, R. S; Wartofsky, M. W., Eds.; D. Reidel Publishing: Dordrecht, 1973.

Arthur, W. B. *The Logic of Invention*. Santa Fe Institute Working Paper 2005-12-045; Santa Fe Institute: Santa Fe, NM, 2005.

Gehring, W. J. The genetic control of eye development and its implications for the evolution of the various eye-types. *Int J Dev Biol* 2002, 46, 65–73.

第08章　技术进化所引发的经济进化

Ashton, T. S., *The Industrial Revolution*, Oxford University Press, New York,1968.

Bartrip, P. W. J. and S. B. Burman, *The Wounded Soldiers of Industry*, Clarendon Press, Oxford, UK, 1983.

Chase, M., *Early Trade Unionism*, Ashgate, Aldershot, UK, 2000.

Fraser, W. H., *A History of British Trade Unionism 1700–1998*, Macmillan,London, 1999.

Landes, David, *The Unbound Prometheus*, Cambridge University Press, Cambridge, UK, 1969.

Mokyr, Joel, *The Lever of Riches*, Oxford University Press, New York, 1990.

Rapp, Friedrich, in Paul. T. Durbin, ed., *Philosophy of Technology*, Kluwer Academic Publishers, Norwell, MA, 1989.

Rose, M. E., “Social Change and the Industrial Revolution,” in *The Economic History of Britain since 1700*, Vol. 1, R. Floud, and D. McCloskey, eds., Cambridge University Press, Cambridge, UK, 1981.

Schumpeter, Joseph, *The Theory of Economic Development*. 1912. Reprinted, Harvard University Press, Cambridge, MA, 1934.

第09章　复杂性的进化

Bonner, J. T. *The Evolution of Complexity*. Princeton: Princeton University Press, 1988.

Constant, E. W. *Origins of the Turbojet Revolution*. Baltimore: Johns Hopkins University Press, 1980.

Darwin C. *From Charles Darwin's Notebooks*, edited by P. H. Barrett et al., 422.Ithaca: Cornell University Press, 1987.

Heron, S. D. *History of the Aircraft Piston Engine*. Detroit: Ethyl Corp., 1961.

Holland, J. *Adaptation in Natural and Artificial Systems*, 2nd ed. Cambridge:MIT Press, 1992.

Holland, J. “Echoing Emergence: Objectives, Rough Definitions, and Speculation for Echo-Class Models.” Mimeograph, University of Michigan, 1993.

Kauffman, S. “The Sciences of Complexity and Origins of Order.” Working Paper 91-04-021, Santa Fe Institute, 1991.

Koza, J. *Genetic Programming*. Cambridge, MA: MIT Press, 1992.

Lieberman, P. *The Biology and Evolution of Human Language*. Cambridge, MA: Harvard University Press, 1984.

Lindgren, K. “Evolutionary Phenomena in Simple Dynamics.” In *Artificial Life II*, edited by C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen. Santa Fe Institute Studies in the Sciences of Complexity, Vol. X, 295–312. Reading, MA: Addison-Wesley, 1991.

McShea, D. “Complexity and Evolution: What Everybody Knows.” *Bio. & Phil.*6 (1991): 303–324.

Morowitz, H. *Beginnings of Cellular Life*. New Haven: Yale University Press, 1992.

Müller, G. B. “Developmental Mechanisms at the Origin of Morphological Novelty: A Side-Effect Hypothesis.” In *Evolutionary Innovations*, edited by Matthew Nitecki, 99–130. Chicago: University of Chicago Press, 1990.

Ray, T. S. “An Approach to the Synthesis of Life.” In *Artificial Life II*, edited by C. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen. Santa Fe Institute Studies in the Sciences of Complexity, Vol. X, 371–408. Reading, MA: Addison-Wesley, 1991.

Stanley, S. M. “An Ecological Theory for the Sudden Origin of

Multicellular Life in the Late Precambrian." *Proc. Nat. Acad. Sci.* 70 (1979): 1486–1489.

Waddington, C. H. "Paradigm for an Evolutionary Process." In *Towards a Theoretical Biology*, edited by C. H. Waddington, Vol. 2, 106–128. New York: Aldine, 1969.

第10章　认知科学

Arthur, W. B. (1994), "Complexity in Economic Theory: Inductive Reasoning and Bounded Rationality," *American Economic Review*, 84 (2), 406–411.

Arthur, W. B., J. H. Holland, B. LeBaron, R. Palmer, and P. Tayler (1997), "Asset Pricing under Endogenous Expectations in an Artificial Stock Market," in W. B. Arthur, S. Durlauf, and D. Lane (eds), *The Economy as an Evolving Complex System II*, Reading, MA. Addison-Wesley.

Clark, A. (1993), *Associative Engines*, Cambridge, MA: MIT Press.

Jaynes, J. (1976), *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, Boston: Houghton Mifflin Company.

Laxness, H. (1935), *Independent People*, New York: Vintage.

Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), 999–1033.

Sargent, T. J. (1993), *Bounded Rationality in Macroeconomics*, New York: Oxford University Press.

结语　复杂的经济需要复杂经济学

P. Anderson, K. J. Arrow, D. Pines, Eds., *The Economy as an Evolving Complex System* (Addison-Wesley, Reading, MA, 1988).

W. B. Arthur, S. N. Durlauf, D. A. Lane, Eds., *The Economy as an Evolving Complex System II* (Addison-Wesley, Reading, MA, 1997).

W. B. Arthur, *Sci. Am.* 262, 92 (1990).

W. B. Arthur, *Increasing Returns and Path Dependence in the*

Economy (Univ. of Michigan Press, Ann Arbor, Ml, 1994).

I have avoided exact definitions of “complexity” and “complex systems.” Technically, the systems I have described are referred to as adaptive nonlinear networks (J. H. Holland's term), and typically if they exhibit certain properties that have to do with the multiplicity of potential patterns or with the coherence or propagation of substructures, they are said to be “complex.” Definitions vary widely.

A. Marshall, *Principles of Economics* (Macmillan, London, ed. 8, 1920), p. 459.

E. Helpman and P. R. Krugman, *Market Structure and Foreign Trade* (MIT Press, Cambridge, MA, 1985).

W. B. Arthur, *Econ.* J. 99, 116 (1989).

W. B. Arthur, *Math. Soc. Sci.* 19, 235 (1990); P. R. Krugman, *J. Polit. Econ.* 99,483 (1991); in (2), p. 239; *Geography and Trade* (MIT Press, Cambridge, MA, 1991).

S. N. Durlauf, in (2), p. 81; *J. Econ. Growth* 1, 75 (1996).

W. B. Arthur, *Am. Econ. Rev*. 84, 406 (1994).

J. Casti, *Complexity* 1 (no. 5), 7 (1995/96); N. Johnson *et al., Physica A* 256, 230(1998); D. Challet and Y. -C. Zhang, ibid. 246, 407 (1997); ibid. 256, 514 (1998).

R. E. Lucas, *Econometrica* 46, 1429 (1978).

W. B. Arthur, J. H. Holland, B. LeBaron, R. Palmer, Paul Tayler, in (2), p. 15; W.B. Arthur, *Complexity* 1 (no. 1), 20 (1995).

See K. Lindgren's classic paper in *Artificial Life II,* C. G. Langton, C. Taylor, J. D. Farmer, S. Rasmussen, Eds. (Addison-Wesley, Reading, MA, 1991), p. 295; H. P. Young, *Econometrica* 61, 57 (1993); L. E. Blume, in (2), p. 425; B. A. Huberman and N. S. Glance, *Proc. Natl. Acad. Sci. U. S. A*. 90, 7716 (1993).

R. Marimon, E. McGrattan, T. J. Sargent, *J. Econ. Dyn. Control* 14, 329 (1990);M. Shubik, in (2), p. 263; W. A. Brock and P. de Lima, in *Handbook of Statistic 12: Finance*, G. S. Maddala, H. Rao, H. Vinod,

Eds. (North-Holland, Amsterdam, 1995).

T. J. Sargent, *Bounded Rationality in Macroeconomics* (Clarendon Press, Oxford, 1993); D. A. Lane and R. Maxfield, in (2), p. 169; V. M. Darley and S. A. Kauffman,in (2), p. 45.

D. C. North, in (2), p. 223.

Y. M. loannides, in (2), p. 129; A. Kirman, in (2), p. 491; L. Tesfatsion, in (2), p.533.

P. Bak, K. Chen, J. Scheinkman, M. Woodford, *Ric. Econ.* 47, 3 (1993); A. Leijonhufvud, in (2), p. 321.

R. Axelrod, *Am. Pol. Sci. Rev.* 80, 1095 (1986); K. Kollman, J. H. Miller, S. E. Page, in (2), p. 461.

I thank J. Casti and S. Durlauf for comments on this article.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读APP：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当做阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”APP。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读APP玩转指南

湛庐阅读APP结构图：



三步玩转湛庐阅读APP：



使用APP扫一扫功能，
遇见书里书外更大的世界！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！



快速了解本书内容，
湛庐千册图书一键购买！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

延伸阅读

《技术的本质》

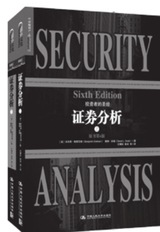
- 复杂性科学的奠基者、著名的技术思想家、“熊彼特”奖的得主布莱恩·阿瑟的经典作品。
- 谷歌董事长埃里克·施密特、硅谷精神教父凯文·凯利、汪丁丁、段永朝、陈劲、包光国等联袂推荐。
- 技术理论体系的先河之作，对科学、技术、经济三者关系提出了全新见解，一次打开“技术黑箱”的尝试性创新探索。





《证券分析》（原书第6版）

- 一本投资者的圣经，自1934年出版以来，历经80年畅销不衰。
- 被誉为“现代证券分析之父”“华尔街教父”、价值投资理论奠基人的本杰明·格雷厄姆的不朽巨著，在投资界享有不可动摇的地位。
- 本书新增了巴菲特的亲笔序言，以及数十位当今华尔街的知名基金经理人所撰写的导读内容。



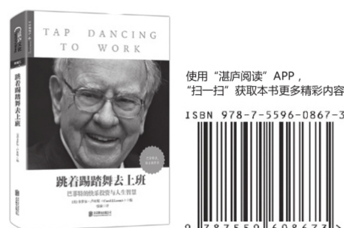


《跳着踢踏舞去上班》

◎　这是一本关于全球著名的投资商巴菲特的故事。

◎　由巴菲特最亲密的朋友、商业新闻界领先大奖“杰拉尔德·洛布奖”“亨利·卢斯奖”的获得者卡萝尔·卢米斯亲自操刀。

◎　本书异常巧妙地将很有价值的文章集结成册，能够清晰地看出巴菲特在商业领域的人生轨迹。

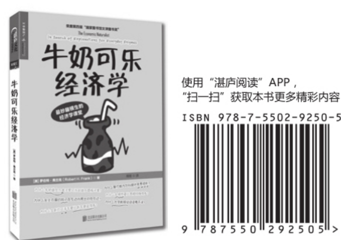


《牛奶可乐经济学》

◎　风靡全球的经典畅销书，作者以专业的角度来阐释生活中的博物经济学理念。

◎　致力于把经济学从数学中解放出来，让深奥的经济学走进普通大众的日常生活中，去解释生活。

◎　本书丰富的案例将加深你对经济学原理的理解。



(1)　Paul Lewis. Emergent properties in the work of Friedrich Hayek，*Journal of Economic Behavior and Organization* vol. 82，2012，pp. 368-378.

(2)　F. A. Hayek. Freedom, reason and tradition, *Ethics* vol. LXVIII，no. 4，1958，pp. 229-245.

(3)　F. A. Hayek. Rules, perception and intelligibility，*Proceedings of the British Academy* vol. XLVIII，1963，pp. 321-344；F. A. Hayek. The theory of complex phenomena，in M. Bunge，ed.，*The Critical Approach to Science and Philosophy: Essays in Honor of K. R. Popper*, 1964，pp. 332–349.

(4)　G. R. Steele. Hayek's sensory order, *Theory and Psychology* vol. 12，no. 3，2002，pp. 125-147.

(5) F. A. Hayek. *New Studies in Philosophy, Politics, Economics and the History of Ideas*, 1978.

(6) H. P. Young. The possible and the impossible in multi-agent learning, *Artificial Intelligence* vol. 171，2007, pp. 429-433；D. P. Foster and H. P. Young. On the impossibility of predicting the behavior of rational agents，*Proceedings of the National Academy of Sciences*，vol. 98，issue 222，2001，pp. 12848-12853.

(7) Ronald Heiner. The origin of predictable behavior, *American Economic Review*，1983，1985.

(8) Alfred North Whitehead. *Modes of Thought*，Lecture Two“expression”.

(9) Gerald Edelman. *Second Nature: Brain Science and Human Knowledge*，2006.

(10) Carl Jung. *Man and His Symbols*，1960.

(11) Roy Wagner. *Symbols that Stands for themselves*，1986.

(12) *The Collected Works of Carl Jung Vol. IX: The Archetypes and the Collective Unconscious*，Second Edition.

(13) Ralph Adolphs. The social brain: the neural basis of social knowledge，*Annual Review of Psychology* Vol.60，2009，pp. 693-716；Tania Singer. The past, present and future of social neuroscience，*NeuroImage* Vol. 61，2012，pp. 437-449.

(14) Agnes Heller，*The Three Logics of Modernity and the Double Bind of the Modern Imagination*, 2000.

(15) Michael Polanyi. *The Tacit Dimension*，1966.

(16) 由湛庐文化策划，约翰·霍兰德所著的经典之作《涌现》（*Emergence: From Chaos to Order*）即将出版。——编者注

(17) 该书构建了关于技术的理论体系，阐明了技术的本质及其进化机制，是技术思想领域的开创性作品。中文版已由湛庐文化策划、浙江人民出版社出版。——编者注

(18) 斯科特·佩奇，美国密歇根大学教授，著有*The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* 及*Model Thinker*，中文版暂定名分别为《差异》和《模型思维》，由湛庐文化策划，即将出版。——编者注

湛庐文化 Cheers Publishing

与思想有关 a business

财富汇

经典版

技术的本质

[美] 布莱恩·阿瑟 ◎ 著
曹东溟 王健 ◎ 译



浙江人民出版社

经典版

The Nature of Technology

What It Is and How It Evolves

布莱恩 · 阿瑟 经典名作

技术的本质

技术是什么，它是如何进化的

圣塔菲研究所元老
斯坦福大学经济学教授
首屈一指的技术思想家
"拉格朗日奖" "熊彼特奖"得主

[美] 布莱恩·阿瑟（W. Brian Arthur）◎著
曹东溟 王健 ◎译



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

技术的本质

技术是什么，它是如何进化的

[美] 布莱恩・阿瑟（W.Brian Arthur）◎著
曹东溟 王健◎译



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2018年6月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有•侵权必究

书名：技术的本质

著者：［美］布莱恩•阿瑟

电子书定价：71.99元

The Nature of Technology: What It Is and How It Evolves

Copyright ©2009 by W. Brian Arthur

赞誉

“技术”这一词汇，或可列为高度流行的日常用语之一。但遗憾的是，不少人将其理解为“工具”或者“技艺”。换个生活化的说法，在很多人眼里，技术是“死”的，但在阿瑟的眼里，技术是“活生生”的，它有自己的“进化”方向，也有自己的“行事”逻辑，甚至技术自身“正在变为生物”。技术并非割裂人与自然的利刃，而是亲近自然、厚爱生命的新物种。让我们透过阿瑟的睿智之眼，领略技术的本质吧！

段永朝

苇草智酷创始合伙人，财讯传媒集团首席战略官

技术是实现人类目的的重要手段，技术的进化对于技术创新政策与管理具有重大的影响。《技术的本质》一书清晰地阐明了技术的定义，睿智地提出了技术自循环的进化律，因此，技术将对人类产生更大的影响，其自身也将获得极大的发展。在技术的进化过程中，技术将如何保持简约并减少“暴力”的产生，这是人类对技术及其进化必须把握的准则。我非常赞同作者的观点，并希望技术给人类带来更多的福祉和良善。

陈劲

清华大学教授，清华大学技术创新研究中心主任

布莱恩·阿瑟关于技术本质的独到见解会启迪所有人，不论是技术的批评者、支持者，还是那些困惑不解的人。

凯文·凯利

《连线》创始主编，畅销书《失控》《科技想要什么》作者

我们的Java，就是根据布莱恩·阿瑟的思想开发的。

埃里克·施密特

谷歌公司前董事长

布莱恩·阿瑟对技术如何发展及其进化过程的分析使我不禁联想到欧几里得几何学，它清晰、简练，而且看起来不证自明，历经多年终于被一位大师表述出来。《技术的本质》是一部开创性的、激动人心的著

作，它极大地丰富了已有的商业、工程以及社会科学的内涵。

理查德·罗兹

《原子弹秘史》作者，普利策奖得主

《技术的本质》是自熊彼特以来关于技术与经济的最重要的一本书。阿瑟通过明晰、易懂的行文，引人入胜的例子，描述着在从石器到iPod的进化过程中，技术怎样“创造着它自己”。这是一部值得被广泛阅读的，具有深入、持久的重要性的著作。读完本书，你将会以全新的方式思考技术。

埃里克·拜因霍克

《财富的起源》作者

每天，在硅谷游弋的成千上万美元都是基于布莱恩·阿瑟的观点。

约翰·史立·布朗

帕洛阿尔托研究中心（PARC）前主席

推荐序一

路径依赖性：人口、经济、技术

汪丁丁

北京大学国家发展研究院教授

学者的头脑，哈耶克把它分为两种，模糊型的和清晰型的。稍后，他补充了一个脚注，称在写《头脑的两种类型》这篇随笔时，他未听说过伯林对学者的划分——只知道一件大事的刺猬和知道许多小事的狐狸。哈耶克自认是一只刺猬。阿瑟也是一只刺猬，多年来，他跨越许多学科追踪研究的唯一重要的课题，可称为“路径依赖性”。

阿瑟是1946年出生的，现在他被称为经济学家，而且在37岁时就成为斯坦福大学最年轻的经济学教授。这些都是事实，但不是全部事实。我清楚地记得在阿瑟1994年出版的《收益递增与经济中的路径依赖性》一书开篇读到这样一则往事：阿瑟在加州理工学院做研究时，发现了经济生活中存在强烈的收益递增性并写文章论述他的发现（我读研究生时也读他的这些文章）。那时他在斯坦福大学粮食研究所任职，可能还担任生物系主任，他与斯坦福大学经济系的两位核心人物共进午餐（听上去是“求职午餐”），在他讲述了自己的“收益递增经济学”之后，一位经济学教授委婉地告诉他世界上没有收益递增这回事，另一位教授更坦率，这位教授可能是当时的系主任，他告诉阿瑟先生，即便有收益递增这回事，我们也不能承认它。这则故事，赫然写在阿瑟著作的开篇。于是这部作品立即入选我的个人藏书——今天，我更乐意收藏电子版。

阿瑟1999年接受“领导力对话”采访时也回忆了这段“痛苦如地狱”的经历，他的描述是：在斯坦福大学的前10年，他发表了许多论文并担任了系主任，然后，他用10年时间试图发表一篇收益递增论文，却因此而离开了斯坦福大学。鼓舞他坚持探索的，是斯坦福大学校园最受爱戴的诺贝尔经济学奖得主阿罗。阿瑟说，阿罗帮助他获得了1987年古根海姆奖学金，并引荐他去圣塔菲研究所任职。又据阿瑟1999年回忆，因新古典增长理论而获得诺贝尔经济学奖的MIT经济学家索罗特意提醒圣塔菲研究所的主持人柯文，说他正在犯一个最严重的错误，因为阿瑟是无名之辈。

阿罗始终为阿瑟的收益递增经济学大肆鼓吹，同样深受阿瑟这一思想影响的，是因新制度经济学研究而获得1993年诺贝尔经济学奖的经济史学家诺思。我在香港大学教书时，于港大书店翻阅诺思1990年出版的《制度、制度变迁与经济绩效》一书时，印象最深的就是他运用阿瑟的收益递增观念于制度变迁的研究。我认为制度在各国，尤其是在中国这样从未中断悠久历史的国家，路径依赖的性质极其强烈。从那时起，阿瑟成为我关注的西方学术核心人物之一。阿瑟的往事永远提醒我，任何主流，包括经济学主流，都不可避免地压迫和排斥人类的独立精神和自由思想。为写这篇中译本序言，我检索了网上关于阿瑟的报道和文章，我发现，那些令人不快的往事完全消失了。这些往事未必是被斯坦福大学别有用心地花钱“遮蔽”了，很可能是因为网络社会的记忆原本就很短暂。

现在，我可以谈正题了。路径依赖性（path dependency），在制度经济学获得诺贝尔奖的那段时期，大约是1985—1995年这十年，对我们这些热衷于中国经济和政治体制改革的学者而言，真是一个最诱人的观念。例如，张五常在诺思得奖时对香港记者大呼“走宝”（即自家的宝贝被人家拿走了）。因为，据说，诺思当年曾在华盛顿西雅图校区听张五常的新制度经济学课程，这相当于师从五常呀。好事的记者于是去问诺思怎样评论张五常的“走宝”慨叹。诺思哈哈大笑，他的评论是，五常言之有理，可是他并未坚持这项研究。读者懂得啦？五常教授20世纪70年代赴香港大学筹建经济系，1982年在芝加哥大学核心期刊《法与经济学》杂志发表了《企业的契约实质》（我评价为他毕生的登峰造极之作），此后，他的注意力转向中国社会制度变迁，再也无暇他顾。

路径依赖性，阿瑟的论述，诺思的论述，以及多年前我的论述，可概括为这样一项平凡的陈述：人的行为依赖于他们过去的全部行为。注意，是“依赖”而不是“由此被决定”，也不是“完全不依赖”。阿瑟早年研究人口学问题，20世纪70年代至20世纪80年代他发表的论文主要是人口学的。不过，他自幼最喜欢数学和工程学，在爱尔兰的少年时代，他偶然选择了电子工程专业，那时他不过17岁——“年轻得有些荒唐”。后来，可能是他在加州理工学院（我认为很可能是北美唯一一所最优秀的理论学院）时期，专注于收益递增现象的研究。直到20世纪90年代主持圣塔菲研究所的“复杂性”课题组，自此以后，他主要研究经济生活中的收益递增现象。

技术，阿瑟指出，不是科学的副产品，而是或许恰好相反，科学是

技术的副产品。古希腊人很早就懂得这一原理。亚里士多德说过，理论家的工作在于冥想，他们的模型是恒星系统，具有“永恒”这一基本性质。技艺是实践者的工作，是一种关于偶然性的艺术，探求永恒原理的哲学家，不愿为也。两千年之后，技术仍是卑贱的实践者的工作（例如米开朗琪罗的工作），却引发了近代科学。

阿瑟继续考证，技术总是由一些基本的功能模块组合而成的。最初的石器，打磨为两类，锋利的和有孔的，与手柄组合而成复合工具，例如“飞去来器”，例如“耜”与“耒”，例如“眼镜”。凡技术发明者，首重适用性和便利性，发明专利所谓“实用新型”。这两大性质要求使用新技术的人群将以往行为与新技术相合。如果你从微软视窗系统转入苹果系统，你会有很多这样的体会，多年之后，你试着适应微软系统，又有很多这样的体会。我们的身体（包括脑内的神经元网络）可以记住我们的行为，并因记忆而有了行为的积累效应——贝克尔称为“人力资本”。在夏威夷的东西方中心人口研究所求学时，一位人口统计学家告诉我，观察人们早餐时吃的是哪一国的食物，可准确判断这些人来自哪一族群。她说，早餐习惯是最难以改变的，因为胃口或口味是“永恒的”。

诺思有几篇论文阐述制度的收益递增效应。他指出，规模越大的政府总是追求更大规模，权力越大的人倾向于追求更大权力，成功的制度有复制自身的冲动，直到社会被锁死于早已僵化但曾经成功的制度陷阱之内。他还找到了不少消亡的人类社会，作为“锁死”效应的例证。诺思的警告格外触动我们这些中国学者，因为历史太悠久而且太难以割舍，所以我们不能放弃传统，但我们必须改造传统，否则中国就可能消亡。

于是，技术的本质，与制度的本质类似，因有强烈的路径依赖性而常将人类“锁入”既有的技术路径或制度路径。锁入，于是可能锁死。当社会被制度路径锁死时，社会消亡。当企业被技术路径锁死时，企业淘汰。现在，读者可以翻阅阿瑟的这部作品了。

推荐序二

打开“技术黑箱”的一个新尝试

包国光
东北大学哲学系教授

美国学者布莱恩·阿瑟的著作《技术的本质》是一次打开“技术黑箱”的尝试性的创新探索，对我国的技术哲学和技术创新研究将具有一定的参考借鉴意义。

布莱恩·阿瑟是研究经济的学者，他在研究报酬收益理论时，发现不能回避技术。他给自己提出相互关联的两个问题：技术是什么，以及它是如何进化的。从西方哲学的角度来看，这相当于在追问“技术的本质”和“技术的实存”。《技术的本质》这部书就是对上述两个重要的“技术哲学问题”的分析解答，作者试图建立一个“关于技术的理论”。

布莱恩发现，当代人对技术很熟悉，而对技术整体又很生疏，“我们知道单个技术（individual technology）的历史以及它们是如何生成的细节；我们可以对设计过程进行分析；关于经济因素如何影响技术的设计、技术被接受的过程，以及技术是怎样在经济中扩散，我们已经有了很精彩的研究；我们对社会如何型塑技术以及技术如何型塑社会进行了细致的分析；我们深思技术的意义，追问技术到底是否决定人类的历史。但是关于‘技术’这个词到底是什么意思，我们并没有达成共识。这里还没有一个关于技术如何形成的完整理论，没有关于创新由什么构成的深刻理解，没有关于技术进化的理论……这里缺失的是某个一般性法则（principles），它可以赋予主体一个逻辑框架，一个有助于填补这鸿沟的框架”。针对这样一种情况，布莱恩从几乎完全空白的状态开始他的研，并主要基于三个基本原理（假设）逐级地构建了他的关于技术的理论：（1）技术（所有的技术）都是某种组合，这意味着任何具体技术都是由当下的部件、集成件或系统组件构建或组合而成的；（2）技术的每个组件自身也是缩微的技术；（3）所有的技术都会利用或开发某种（通常是几种）效应或现象。

布莱恩对技术本质的探讨是从分析技术的构成结构开始的。技术有

其自身的结构。结构首先是指技术是由零部件构成的。技术的最基本结构包含一个用来执行基本功能的主集成和一套支持这一集成的次集成。技术包含的集成块是技术，集成块所包含的次一级的集成块也是技术，次一级集成块包含的再次一级的集成块还是技术。这样的模式不停地重复，直到最基础水平的基本零件为止。换句话说，技术有一个递归性结构。技术包含着技术，直到最基础的水平。这种组合结构一直分解下去，将到达一类不再属于技术的“现象”或“效应”那里。这些现象具有恒定性和可重复性，独立于人类的技术和科学而存在，在人类的技术活动和科学研究活动中显现并被捕捉。布莱恩举例解释了人们对现象的利用：炼制石油是基于气化原油中的物质会在不同的温度下凝结这样一种自然现象；一个下落的锤子则依赖于动量的传输现象；汽车则依赖汽油或柴油燃烧后产生能量这种自然现象。布莱恩这里所说的“现象”，显然不是胡塞尔意义上的“现象”，也不是马克思、黑格尔和康德意义上的“现象”。布莱恩的“现象”是自然现象或“自然效应”，所有技术都建立在对这种自然现象的利用之上。“自然现象是技术赖以产生的必不可少的资源。所有的技术，无论是多么简单或者多么复杂的技术，实际上都是应用了一种或几种现象之后乔装打扮出来的。”

关于技术的进化和创新，布莱恩提出了技术是“自我创生的”（autopoietic）观点。所有技术的产生或使其成为可能，都源自以前的技术。技术是从已有的技术中产生的，是通过组合已有技术而来的。在这个意义上，技术集合（the collective of technology）的新元素的产生或成为可能，正是源于已有的技术集合，结果就是技术创生于技术自身。这样一来，所有技术产生于已有技术，也就是说，已有技术的组合使新技术成为可能。但同时，布莱恩也强调了：“说技术创造了自身并不意味着技术是有意识的，或能以某种阴险的方式利用人类为它们自身的目的服务。技术集合通过人类发明家这个中介实现自身建构，像珊瑚礁通过微小生物自己建构自己一样。假如我们把人类活动总括为一类，并把它看成是给定的，我们就可以从这个意义上说，技术体是自我创生的，它从自身生产出新技术。”从这里可以看出，布莱恩持有“技术自主论”立场。

纵观全书，作者提出了不少新概念和新见解，如递归性组合、现象、域、链接、自创生、珊瑚礁结构等新概念，以及“技术由技术构成”“技术是对现象的编程”“技术发明是需求和现象的链接”“技术进化的自创生”“技术与科学的同源性”“经济随技术而进化”等观点，对我国技术哲学和技术创新的研究具有重要的启发作用。

布莱恩对“技术是什么，它是如何进化的”之回答，以及书中的“现象”的优先论、“技术自主论”、“域”本身的转换论等观点，虽然并非无懈可击，但也因得到理论和事实的论证支持而部分成立，这需要我国技术哲学和技术创新研究者进一步深入探讨。

本书的译者从事技术哲学和技术创新的研究与教学工作多年，熟悉技术哲学和技术创新领域的范式和研究状况。她们对原著中探讨的问题与思想有深刻的理解，基本把握了原著的分析思路和思想精髓。译者对原著中的一些新概念术语也给予了妥帖的转化，使读者能够基本准确地领会作者的思想用意。

前言

技术的追问

人在十几二十岁的时候，常会碰到一些没办法解答的问题，它们可能就此盘踞于心，很久都无法释怀。我是17岁开始接受电子工程本科教育的，我很快就意识到，其实我并没有真正理解我所学的东西的本质，即什么是技术真正的本质。尽管那时我可以得到很高的分数，但我认为那只是因为我的数学还不错。教授们解释道：技术是科学的应用；技术是经济中关于机制和方法的研究；技术是工业过程中的社会知识；技术是工程实践。但是所有这些答案似乎都不能令我十分满意，没有哪个答案触及“技术的本性”（technology-ness）这个层次。因而对我来说，它始终是一个未解之谜。

后来，到了研究生阶段，我转而开始着迷于经济是怎样发展并建构起来的这个问题。对我来说，很明显，经济很大程度上是从技术中产生的。毕竟，在某种意义上，经济不过是通过明智地组织技术来满足我们的需求，故而它也会随技术的进化而进化。但如果真是这样的话，技术是如何产生的呢？它们是从哪里来的呢？经济又是如何引发技术的？准确来讲，技术到底是什么呢？这样一来，我就又回到了老问题上。

其后的很多年，我都没有再过多地思考这个问题。直到20世纪80年代，当我开始研究收益递增理论时，我的注意力才被重新拉回到技术上。技术，是新的技术产品和生产工艺（例如，早期的汽车）通过被应用和被采用而获得改善，之后再获得进一步的应用和采用，在被采用的过程中形成正反馈或收益递增。收益递增向经济学提出了一个问题：假如有两种收益递增的产品（也可以指两项技术）相互竞争的话，领先的那个就有可能进一步领先，并因此主导市场。但是最终赢家却无法确定，这其中会有多种可能性。那么赢家是如何被选择的呢？在我的理论进路中，是允许这种随机事件发生的，它会被内在的、连续的正反馈所放大，可以随时间随机地选择结果。如果我们将其看成某种程度的随机过程，我们就可以分析收益递增的情况。这样一来，思路顿开。

为了寻找合适的例子，我从1981年开始关注具体技术及其产生和发展的过程。这些考察对我的理论建构都很有帮助，但实际上吸引我的并

不是那些直接与收益递增相关的技术，而是在技术呈现之初，那些看起来模模糊糊的状态。我意识到，新技术并不是无中生有地被“发明”出来的，我看到的技术的例子都是从先前已有的技术中被创造（被建构、被聚集、被集成）而来的。换句话说，技术是由其他的技术构成的，技术产生于其他技术的组合（combinations）[1](1)。这个观察结果看起来太简单了，以至于一开始会让人觉得并不特别重要，但是我很快意识到，如果新技术是从已有技术中建构出来的，而且被当作一个整体来看的话，那就意味着技术自己创造了自己！后来，我接触到了弗朗西斯科·瓦雷拉（Francisco Varela）和温贝托·马图拉纳（Humberto Maturana）的自创生系统理论（self-producing systems）。我知道，如果我直接采用“技术是自我创生的（autopoietic），或者自我创造的（self-creating）”这样的阐述，其实可以令读者印象更加深刻。但是在20世纪80年代，我根本不知道瓦雷拉和马图拉纳。当时我能做的只是观察这个自我创生的对象的宇宙，惊讶于这种自创生的结果。

我逐渐意识到，“组合”可能是弄清楚技术的发明与进化的现实机制的关键所在，在此之前，这些想法还没有被技术思想家认真思考过。我在20世纪90年代对一些机制有了一些想法，并在1994年发表过关于结构深化的文章。与此同时，我也对其他理论有了一些模糊的理解。

20世纪90年代，我曾研究了一些其他议题，主要是关于经济中的复杂性和认知的议题。直到2000年，我才又开始回过头系统地思考技术以及技术是如何产生的问题。我慢慢悟出，除了“组合”之外，还有其他原理也在起作用。技术是由部件和零件（集成件和次级集成件）构成的，而集成件自身也是技术。所以技术有一个递归性的（recursive）结构。而且我认识到，每个技术都是建立在某个现象，以及从该现象挖掘出来的某种或几种效应之上的。因此从整体来看，技术是通过捕捉现象并对之加以应用来获得发展的。同时，我也认识到，经济并不太像我接受的教育所暗示的那样，是技术的集装箱，经济是从技术之中产生出来的。经济是从满足我们需要的生产性的方法、法规和组织性安排当中产生出来的，因此经济是从捕捉现象及之后的技术组合过程中发展起来的。

为了深入思考，我一头钻进了斯坦福图书馆。一开始，我需要阅读的资料似乎非常多。但是随着阅读与思考的深入，我又觉得可读的材料实在太少了。这很奇怪，因为关于技术的资料应该和经济、法律法规之类的资料同样庞大、复杂和有趣。我看到图书馆中有大量关于具体技术的文章、丰富的教科书，特别是关于那些流行的技术，如计算技术和生

物技术的读物。但是关于技术或技术创新的本质，以及其后续进化的相关文献却很少。这些资料里有工程师和法国哲学家关于技术的沉思，有关于技术的采用与扩散的研究，有关于社会如何影响技术以及技术如何影响社会的理论，还有关于技术是如何被设计、如何发展的观察，但是当我想要追问得更深刻一些，想讨论技术背后的原理，以及建构技术并决定其方式与过程的通用逻辑的时候，却没能发现更进一步的论述。因此我假定，这可能意味着我们还没有一个关于技术的完整理论。

在这本书中，我将讨论所有我能找到的关于技术思考的文献，它们涵盖了来自哲学家、工程师、社会科学家和历史学家等一小撮思想家的相关论述。所有这些讨论都很有帮助，其中最有用的是历史学家们对一个个具体技术形成的细节以及细致的案例研究。[2]开始，我搞不懂为什么在所有这些思考技术的人里，历史学家在技术和创新的方法和本质方面表达最多，后来我明白了，世界上有更多东西是从技术、而不是战争和条约中涌现出来的，而历史学家当然关注世界是如何形成的，因此就会对技术是怎样形成的更感兴趣。

本书讨论的议题是，技术是什么，它是如何进化的。这主要是在我的两场学术报告的基础上完成的：一个是1988年在圣塔菲研究所“斯塔尼斯拉夫·马尔钦·乌拉姆（Stanislaw Marcin Ulam）纪念演讲”上关于“数字化与经济”的一个讲座；另一个是2000年在戈尔韦的爱尔兰国立大学的“凯恩斯论坛（Cairnes Lectures）”上关于“高科技与经济”的那场报告。本书内容大部分基于上述两个报告，但主要来自“凯恩斯论坛”的报告。

在撰写本书时，我不得不作一些决定。其中之一是，我决定用平实的语言来写这本书（或者我希望它是直白的）。从职业和实质来讲，我是理论家，所以我必须承认这么做会有些顾虑。为大众写作一本关于严肃理论的书在100多年前是很普通的事，但是今天这么做的话，人们很可能会认为你不够严肃。当然，在我最熟悉的经济学和工程学领域里的人要表现得卓尔不群，那就要用专业、晦涩的术语来表达自己的理论。

当然，出于多种理由，要写一本“既严谨又能满足大众阅读需求”的书的想法最终还是胜出了。主要理由是：首先，单纯的诚实性需要。由于研究对象在此之前没有被详细地思考过，所以还不需要晦涩的专业术语的介入；其次，我认为技术太重要了，因此不能为专家所独有，而是需要普罗大众的共同参与；最后，很重要的是，我要激起公众广泛关注这样一个异常美丽的主题的兴趣，关注我所坚信的那个技术背后必定拥

有的某种自然逻辑。

最早，我就发现在技术研究领域“用词”本身是个问题，技术中的许多词汇都被滥用了，比如，“技术”（technology）这个词本身，以及“创新”（innovation）、“技艺”（technique）。它们的内涵往往既相互重叠，又经常相互矛盾。仅“技术”一项就至少有半打的主要定义，而且相互之间多有含义上的冲突。另外几个词又常常引起人感性的联想，比如，“发明”常常会使人不禁在头脑中浮现出一个孤独的发明家独自与“或然性”作战的情景。这种情景会使人误会新技术是来自天才们紧蹙的眉间，而不是衍生于此前的技术。我开始意识到，许多技术思考的困难可能恰恰源自用词。随着研究的深入，我发现自己做的是有点类似于数学家的工作：首先需要准确界定术语，然后由此逻辑地导出结果和性质。结果将如读者所见，我需要不断地（而且必要地）关注词汇以及它们在技术中的应用，必要时还要引进几个新的术语。我希望能尽量避免这种情况，但是为了讨论需要，还是额外引入了几个术语。

另外，尽管我一直坚持认为在很窄的案例范围内进行论述可能会更方便，但我还是要从更大的范围内选定案例。一位优秀的出版人曾经建议我就用铅笔做例子，但是我认为既然对技术而言，存在着一个既适用于计算机算法，也适用于啤酒酿造，既适用于发电站，也适用于铅笔、掌上游戏机和DNA测序技术的通用逻辑，那么，案例就应该覆盖所有类型的技术。当然，为了使我的论述更加明晰且省去太多不必要的解释，我会选择读者较为熟悉的技术。

最后，我还要对这本书不想做的事说上两句。首先，它不是对未来社会和环境所作的技术承诺或者威胁，这些论题都很重要，但并不是我在这里要讨论的内容；它不是关于具体技术，不是关于即将出炉的某个新技术，也不是关于某个工程过程的机械论的概述，那些都已经被广泛谈论过了；同时，它也不是关于人类创造技术的讨论。尽管在技术创造过程中的每一步都有人的参与，但是我的注意力将会集中在驱动这个过程的逻辑上，而不是放在卷入其中的人身上，我一开始就决定只讨论直接相关的主题。还有另外几个有价值的主题我只是一笔带过，比如：发明社会学、技术的采用和扩散、成本推动和需求拉动理论、制度和学术团体的作用，还有技术史。所有这些理论都很重要，但是在这本书里都没有着重提及。

尽管本书一定会涉及关于技术的相关文献，但本书并未对其进行回顾。我常常想起刘易斯（Lewis）和克拉克（Clark）的探险[2]：他们每

次都从最熟悉的地方开始探险之旅，迅速到达一个新的地方，偶然会回到以前曾被别人占领过的地方。我的这次探险也不例外。我们会遇到一些从前的旅人。在这个领域，海德格尔留下过足迹，而熊彼特的足迹到处都是。此外，还有许多学者对该领域进行了研究，本书或对他们的研究成果有所疏漏，在此向这些探险者一并致歉。

最后一个免责声明是，读者不要因为我写了一本关于技术的书，就认为我对技术情有独钟。脑瘤科医生可能会写关于癌症的书，但是那并不意味着他们希望它发生在某个具体的人身上。我对技术以及技术后果都持怀疑的态度。但我也必须承认，我对科学怀有激情并着迷于技术的魔力，而且我也得承认我热爱飞行器，也热爱老式无线电。

想要对技术的本质一探究竟吗?



扫码获取“湛庐阅读”APP，
搜索“技术的本质”，
查看本书彩蛋！



什么是彩蛋

彩蛋是湛庐图书策划人为你准备的更多惊喜，一般包括①测试题及答案②参考文献及注释③延伸阅读、相关视频等，记得“扫一扫”领取。

目录

赞誉
推荐序一　路径依赖性：人口、经济、技术
推荐序二　打开“技术黑箱”的一个新尝试
前言　技术的追问
第1章　问题

技术给我们带来了舒适的生活和无尽的财富，也成就了经济的繁荣。一句话，我们的世界因技术而改变。但是，技术的本质究竟是什么呢？它从何而来，又是如何进化的呢？

缺失了本质的技术
技术的进化
组合进化
本书的主题
第2章　组合与结构

新技术都是在现有技术的基础上发展起来的，现有技术又来源于先前的技术。将技术进行功能性分组，可以大大简化设计过程，这是技术“模块化”的首要原因。技术的“组合”和“递归”特征，将彻底改变我们对技术本质的认识。

技术结构的形成
为什么要模块化
递归性及其重要性
第3章　现象

无论是简单还是复杂的技术，都是在应用一种或几种现象之后乔装打扮出来的。技术就是那些被捕获并使用的现象，是对现象有目的的编程。我们一直以为技术是科学的应用，但实际上却是技术引领着科学的发展。

技术的本质
有目的的系统
捕捉现象
技术与科学
第4章　域

为了共享现象族群和共同目标，或者为了分享同一个理论，个体技

术就会聚集成群。这种集群就形成了“域”。工程设计是从选择某个域开始的，这个自动和下意识的选择过程叫作“域定”。设计工作就像是用某种语言所进行的写作或表达。

域定
设计就如同语言表达
参与的世界
第5章　工程和对应的解决方案

几乎所有的设计都是某个已知技术的新版本，只有在必要的情况下，才需要一项全新的设计。工程师在寻找解决方案的过程中，把适宜的构件选择出来，让它们组合在一起工作。设计就是选择，组合只不过是工程的副产品。工程的解决方案，又成为发展新技术的新构件。

标准工程
解决问题的工程
组合与解决方案
未来的技术构件
第6章　技术的起源

新技术可以是根据某个目的或需要发现一个可以实现的原理，也可以从某一新现象出发，找到如何使用这种现象的办法。原理可以借用，也可以是先前概念的组合，或者由理论而来。只有将概念转化为现实，一项新技术才真正诞生。科学和数学中的原创，与技术没什么两样，因为它们同属“目的性系统”。

什么样的技术才算新技术
找到一个基本原理
概念的物化
基于现象的发明
什么是发明的核心
因果性金字塔
科学与数学中的发明
发明与新的构件
第7章　结构深化

技术一旦走上发展之路，各种各样的版本就会随之出现。通过“内部替换”，开发人员可以用更好的部件（子技术）更换某一形成阻碍的部件。开发人员还可以通过寻找更好的部件或材料，或者加入新组件进行结构深化。旧设计和旧原理一经锁定，就会产生新用途。

内部替换
结构深化
锁定与适应性延伸
第8章　颠覆性改变与重新域定

域并不是若干单个技术的简单相加，它们是连贯的整体，对经济的影响也更大。任何新域，都产生于一个已存在的域——母域，而且参与者开始很少能意识到会发生“颠覆性改变”。但随着理解的深化和实践的固化，新域会慢慢忘记它的母域而横空出世，甚至会极大地提升国家竞争力。

域是如何进化的
经济的重新域定
经济中的时间
创新与国家竞争力
第9章　进化机制

组合是新技术的潜在来源。组合的威力，在于它的指数级增长。如果新技术会带来更多的新技术，那么一旦元素数目超过一定阈值，可能的组合数就会爆炸性增长。此外，机会利基也在呼唤着新技术。技术就如同生命体一样，它的进化与生物进化也没什么本质差异。

组合
机会利基
核心机制
关于进化的一个思想实验
技术进化与生物进化的比较
第10章　技术进化所引发的经济进化

众多的技术集合在一起，创造了一种我们称之为“经济”的东西。经济从它的技术中浮现，不断从它的技术中创造自己，并且决定哪种新技术将会进入其中。每一个以新技术形式体现的解决方案，都会带来新的问题，这些问题又迫切需要进一步得到解决。经济是技术的一种表达，并随这些技术的进化而进化。

经济就是技术的一种表达
结构性变化
解决方案带来的新问题[7]
第11章　我们的立场是什么

随着基因组研究和纳米技术的发展，生物正在变成技术。与此同

时，从技术进化的角度看，技术也正在变为生物。两者已经开始相互接近并纠缠在一起了。我们需要和自然融为一体。如果技术将我们与自然分离，它带给我们的就是死亡。如果技术加强了我们和自然的联系，那就是它对生命和人性的厚爱。

技术具有生物属性，反之亦然
繁衍性经济
纯粹秩序与混质活力
我们应该怎样看待技术
注释
译者后记

THE

NATURE

OF

TECHNOLOGY

第1章 问题

技术给我们带来了舒适的生活和无尽的财富，也成就了经济的繁荣。一句话，我们的世界因技术而改变。但是，技术的本质究竟是什么呢？它从何而来，又是如何进化的呢？

关于技术，我百感交集。使用技术时，我总是觉得理所当然；我享受技术带给我的便利，但偶尔也会因技术而产生出某种挫败感；我无意识里对技术怀有疑问，时常暗暗追问它到底对我们的生活做了什么；我更常常沉醉于技术的奇迹——这个由人类创造的奇迹。匹兹堡大学的研究者们开发出了一种技术：在猴子的大脑中植入微小的电极，然后让猴子通过这些电极来控制一只机械臂。结果是，猴子不需要任何诸如颤动、眨眼或者其他微小的动作，而仅仅通过思维（thought）就可以指挥动作的发生。

这里的技术实际上并非特别复杂，它的构成不外乎是电子科学和机器人领域中的标准元器件，包括用来探查猴子大脑信号的一些电路，用来将信号转化成机械运动的一些处理器和机械驱动器，再加上用来将触觉反馈到猴子大脑的一部分电路。[1]这个实验的真正价值在于理解“意向”（intend）这个动作的神经回路，通过正确的“电路搭接”，就可以让猴子运用这样的回路去移动机械臂了。这种技术对残障人士具有显而易见的好处，但更令人感到神奇的是，当我们把一些电路和机械联动装置（说到底，就是一些硅晶片、导线、金属丝以及小齿轮等）连起来之后，思维，仅仅是思维就可以引起机械运动！

还有许多事令我感到不可思议。把合金片和矿物燃料放到一起，我们就可以以音速直冲苍穹；把原子核自旋产生出来的信号加以组织，就可以显现出我们大脑中的神经回路图像；我们甚至可以对生物对象（例如，酶）进行操作，剪断其DNA中微小的分子片段，再将它们粘进细菌细胞之中。所有这些，在两三百年前是无法想象的，现在我们则拥有了这些力量。我们是如何获得这些力量的？对于我来说，这是个奇迹。

大多数人并没有停下来并深入思考技术。我们发现技术很有用，而技术慢慢地变成了我们世界的背景。对我来说，另一个奇迹是，正是这个背景在创造着我们的世界。我们赖以栖息的家园实际上是由技术创造的。假如某天早上醒来后，你发现由于某种神奇的魔法，过去600年来的技术统统消失了：你的抽水马桶、炉灶、电脑、汽车统统不见了，随之消失的还有钢筋水泥的建筑、大规模生产方式、公共卫生系统、蒸汽机、现代农业、股份公司，以及印刷机，你就会发现，我们的现代世界也随之消失了。如果这奇怪的事情真的发生了，我们的思想、文化可能还在，我们的子孙、配偶也在。当然我们将仍然拥有技术，我们会有水磨、铸工工场、牛车，还会有粗亚麻布、带帽斗篷，以及复杂的教堂建

筑技术。只不过，我们也将会再次经历中世纪。

是技术将我们与中世纪分离的，的确，是技术将我们与我们拥有了5万年甚至更久的那种生活方式分开了。技术无可比拟地创造了我们的世界，它创造了我们的财富，我们的经济，还有我们的存在方式。

那么，技术，这个如此重要的东西究竟是什么呢？在本质上，在最深切的本质上，技术是什么？它从何而来？它又是如何进化的呢？

这些都是我要在本书中追问的。

> 技术思想前沿
>
> 技术的循环：技术总是进行这样的循环，为解决老问题去采用新技术，新技术又引起新问题，新问题的解决又要诉诸更新的技术。

或许我们只需要简单地接受技术，而不必卷进技术背后那些深层问题之中。但是我认为，实际上我是强烈地认为，理解技术是什么以及它是如何形成的，非常重要。这不仅因为技术创造了我们大部分的世界，而且还因为，无论我们是否注意到，在我们这个历史阶段，技术已经令人类感到压迫，感到困扰了。当然，技术的确给我们带来了繁荣。在仅仅两三百年的时间里，技术的发展使一些在以前可能会夭折的孩子现在却得以存活，技术使人类的寿命延长了，使我们比先辈们过得更舒适了，但同时，技术也给我们带来了深切的不安！

这种不安不仅只是来自于恐惧，害怕技术给它所解决的每一个问题带来更多的新问题。这种不安还从更深层的无意识中涌现出来。我们寄希望于技术，我们期待技术能使我们生活得更好，能解决我们的问题，能使我们摆脱困境，能为我们和子孙后代提供想要的未来。然而，作为人类，我们实际上不应该和技术如此紧密地结合，而是应该和其他什么东西融合得更为紧密，那就是自然。在最深的层次上，人的存在应该和自然，和我们最初的环境，以及最初使我们成为人的那些条件相融合。我们熟悉自然，依赖自然，缘于自然是我们300万年的家。我们在骨子里信赖自然！当我们邂逅技术，比如干细胞再生治疗，我们会心怀希望，但我们也立即会问：这样做是否自然？这样一来，我们会被两种巨大的、无意识的力量所左右：一个是我们人类寄托在技术上的最深切的希望；另一个是我们对自然的最深切的信赖。这两种力量就像是地壳板块在漫长又缓慢的碰撞中，不可阻挡地相互挤压、相互融合。

当然，这种碰撞由来已久，但它更强烈地在定义我们所处的时代。技术在加速创造我们这个时代的议题和巨变。我们从一个用机器强化自然的时代（提高行动速度、节省体力、织补衣服）到达了一个用机器来模仿或替代（resemble）自然的时代（基因工程、人工智能、医疗器械身体植入）。随着我们学习、应用这些技术，我们渐渐从应用自然，发展到直接去干预自然。所以这个世纪上演的故事将是，技术所提供的东西与我们感到舒服的东西之间的碰撞。迄今为止，还没有人下结论说，技术的本质和作用方式是简单的，也没有理由认为在本质和作用方式上，技术比经济或者法律简单。但是我们必须注意，技术正在决定我们的未来以及我们会为之焦虑的事情。

这不是一本评价技术好坏的书，市面上有许多书是专门讨论这类问题的。这本书试图去理解这个为我们的世界贡献良多、同时又引起了我们无意识的不安的技术，它本质上到底是什么。

这又把我们带回了同样的问题，什么是技术？它的最深的本质是什么？它的特性和原理是什么？它从哪里来的？它是如何形成的？它又是如何发展的？以及，它是如何进化的？

缺失了本质的技术

一个好的开篇也许应该先这样问：关于技术，我们真正知道些什么？读者可能期望从这里就可以直接得到答案，但那是不可能的。实际上，这几乎是个悖论：我们对技术了解很多，同时又知之甚少。关于一个个具体的技术，我们知道得非常多，但是在一般意义的理解上，我们对技术的了解又很少。我们，或者至少一些人（例如，设计者）应该知道许多使用技术的具体方法、实践程序以及机械装置的特性。我们知道生产计算机的每个步骤，计算机的每个部分，以及每个部分的每个部分。我们确切地知道计算机是如何运行的，甚至知道里面电子的运行轨迹。我们还知道处理器是如何与计算机中其他的元器件相匹配的，以及它是如何与BIOS（基本输入/输出）芯片和中断控制器接口的。我们确切地知道所有这些事情，确切地知道每项技术中存在着什么——因为是我们将它们所有的细节安置到位的。技术实际上是人类经验中最完整的已知部分之一。然而关于它的实质——它的存在的最深的本质，我们却知之甚少。

这种知其然不知其所以然的事情并不少见。大约两个世纪前，在法

国动物学家居维叶的时代，生物学（当时被称为博物学）是一个关于个体物种和比较解剖学，以及它们之间关系的庞大知识体。居维叶在1798年就已经说过："今天，比较解剖学到达了完美的顶点，以至于单看一根骨头，就有可能确定其在生物分类中的'纲'，有时甚至能确定它的'属'。"[2]居维叶只是略有夸大，自然科学家确实拥有详尽的知识，并且他们洞悉动物之间的家族关系。但是他们没有几条可以把握所有知识的原理。他们不太清楚动物是如何而来的；他们没有一个所谓"进化机制"（如果真的存在的话）可以拿来运用；他们也不太清楚动物是否可以在局部对自身进行修改，或者说，他们不知道这种改变是如何发生的。所有这些都要等到那些原理被发现之后才能明了。

面对技术，我们也处于同样的境地。我们知道单个技术（individual technology）的历史以及它们是如何生成的细节；我们可以对设计过程进行分析；关于经济因素如何影响技术的设计、技术被接受的过程，以及技术怎样在经济体中扩散，我们已经有了很精彩的研究；我们对社会如何型塑（shapes）技术以及技术如何型塑社会进行了细致的分析；我们深思技术的意义，追问技术到底是否决定人类的历史。但是关于"技术"这个词到底是什么意思，我们并没有达成共识。这里还没有一个关于技术如何形成的完整理论，没有关于创新由什么构成的深刻理解，没有关于技术进化的理论。还没有一整套总体原理，通过这些原理来给我们的研究主题提供一个逻辑框架，而这个逻辑框架又有助于填补上面所说的那些空白。

换句话说，我们缺失的是一个关于技术的理论——一门关于技术的"学"。

技术思想前沿

关于技术的理论之所以缺失，是因为：

- 技术一直处于科学的阴影之中。
- 那些认真思考技术的人大多数都是社会科学家和哲学家。

我们还没有明显的理由能够解释为什么会这样。但是我非常怀疑，造成这些空白的一个重要理由是，技术一直处于它久负盛名的"姐姐"——科学的阴影之中，这导致我们给予技术的尊重较少，因此对它的研究也较少。我还怀疑这是因为我们觉得技术引起了我们这个世界许多

不和谐事件的发生，在某些无意识层面上，我们觉得技术在智识上是令人反感的——也许不值得深入研究。另一方面，还可能因为我们依稀觉得，既然我们已经创造了技术，我们就已经很了解它了。

还有一个理由，那些最认真思考技术的一般性问题的人大多数都是社会科学家和哲学家。可以理解，他们往往从技术的外部将技术当作独立的对象来看待。如蒸汽机、铁路、贝塞麦炼钢法、发电机，所有这些技术被当作是看不见其内部的箱子，用经济史学家内森·罗森伯格（Nathan Rosenberg）的术语说就是“黑箱”。技术被“黑箱”藏了起来，其内部无法显现。如果我们希望知道的只是技术如何进入经济生活，以及是如何展开的，那么这种从技术外部看技术的方式是足够的。但是如果我们感兴趣的是那些根本性的问题，这样做就不够了。这就像把动物世界看做不同物种的“黑箱”的集合：狐猴、猕猴、斑马、鸭嘴兽等，不厘清它们之间的关系，没有内部解剖学的比较，其实是很难看清它们是如何相互联系，如何起源以及如何进化的。看待技术也是如此。如果我们想知道技术是如何相互联系，如何起源以及如何进化的，我们需要打开它们，去看看它们内部的“解剖学”关系。

公平地说，社会科学家知道技术是由内部组件（components）构成的。在许多情况下，他们也知道内部组件是如何组合在一起使技术得以产生的。而且部分历史学家曾经“打开”过许多技术黑箱，详细探究了这些技术的起源及其随时间的变迁历程。但是这些“内部思考”大多只是关注某项具体的技术，如半导体、雷达、互联网，而不是一般意义上的技术。如果工程师能够一直担当技术的主要思考者的话，事情可能就不一样了，因为他们是自然而然的“内部思考者”。但有一次，当我问著名的技术专家沃尔特·文森蒂（Walter Vincenti），为什么绝少有工程师尝试去奠定他们领域的理论基础，他的回答是：“工程师只喜欢那些他们能解决的问题。”

技术的进化

我想解决的当然是关于技术的比较深层次的问题之一，即技术是如何进化的？或者我应该问：技术是否会进化？因为还没有毫无争议的观点认为，技术的确是进化的。“进化”这个词有两个一般性的含义[3]：一个是某事物渐进的变化，就像芭蕾或者英国情歌的“演变”，我称这种进化为狭义的进化，或者更像“发展”（development）；另一个含义是

指，某类事物的所有对象联结在一起的过程，而其联结纽带也在于它们诞降（descent）自相同的先前对象的集合[3]。这是进化的完整含义，也正是我所谓的进化（evolution）。

对我来说，技术如何进化是技术的核心问题。我为什么这样认为呢？因为如果没有进化，没有某种普遍的关联性，技术看起来就好像是自己独自产生出来，独自改进的。任何技术都一定来自于一些无法解释的心理过程，诸如所谓的“创意”或“跳出原有框架思考”，从而带来新技术，并且独立地发展它。新技术通过进化（如果我们能发现它是如何工作的），以某种精确的方式从以前的技术中“诞生”出来，在此过程中，还需要强大的精神助产士的支持。换句话说，如果我们能够理解进化，我们就能理解那个最神秘的过程：创新。

> 技术思想前沿
>
> “进化”的完整含义：某类事物的所有对象联结在一起的过程，而其联结纽带也在于它们诞降（descent）自相同的先前对象的集合。

技术的进化观点根本不新鲜。达尔文的《物种起源》发表仅仅4年之后，塞缪尔·巴特勒（Samuel Butler）就提出了“机械王国”理论，希望能够解释“机械中的那个相当于动植物王国中自然选择的部分”。他的论文，《机械中的达尔文》，充满了时代激情：

> 没有什么比看到两个蒸汽机之间发生可以繁衍的联姻，让我们这个会痴迷机器的物种更期待的了，而这现在居然成真了。如今机器被用来生产机器了，同时它又变成了以后同类机器的父母。当然，距离机器间的联姻与调情、求爱和婚配看起来还非常遥远。

这当然是一种夸张。然而，如果认真来看这篇文章的话，我就不可避免地会认为，巴特勒当时是在努力将技术塞进某个类似达尔文生物进化论的理论框架，尽管那样可能并不适合。

历史记录清晰地表明，现代某些技术确实是其先前技术的后代。巴特勒之后大约70年，社会学家吉尔菲兰（S. Colum Gilfillan）对船的谱系（从独木舟到帆船，再到当时的蒸汽船）进行了追踪。[4]吉尔菲兰是美国历史和社会学学派的少数成员之一，对19世纪二三十年代的技术和

创新非常感兴趣，而且谙熟关于船的知识。他曾经是芝加哥科学与工业博物馆轮船展厅的负责人。1935年，他从历史细节中追查、研究船壳板、壳板板肋、紧固连结件、龙骨、斜挂大三角帆，以及横帆是如何发明出来的（仅仅描述斜桁帆的起源他就用了4页纸）；从最原始的漂流物到帆船，所有这些是怎样慢慢地演变的；以及以帆船为原型的发明又是怎样演化成现代的蒸汽船的。但是，这还不是整体意义上的进化，只是狭义的进化，即渐变，一种形态上的延续演变（the descent of form）。吉尔菲兰的例子表明：对于某些技术，比如船，我们是可以追查出一个详细的谱系的。

但是，为了得到技术进化的完整理论，我们还需要更多的东西。我们需要找到一个理由去证明所有技术，而不只是某些技术，产生于之前技术，还要找到这种繁衍发生的明确的机制（mechanism）。已有的寻找理由的尝试既少又不成功。大多数的努力，像巴特勒所提供的，都是建设性的意见，而不是理论，而且大多直接将其论证建立在达尔文理论之上。其核心思想是这样的：一项给定的技术，例如铁路机车，在某一特定时间内会有许多变体。这是因为它要达到的目的不同，操作的环境不同（你也可以说，它要适应的“生境”(4)不同），还有不同的设计者提供的设计理念也不同。在这些变种当中，某些变体表现得更好并被选择做进一步的应用和发展，这些变体的细小差异又被传递到未来的设计中去。这样我们就可以跟着达尔文说，“当对个体有利时，通过自然选择，这些小差异会持续地积累，从而导致整个结构更重要的调整。”技术就是这样进化的。

以上论述听起来很有道理，但是很快它就遇到了困难。某些技术，例如激光、喷气机、雷达、计算机快速计算程序以及铁路机车等，在自身刚刚出现时，或至少在即将出现的时候，它们并不是之前技术的样子，这点和新生物种的情况并不一样。喷气机不是内燃机或任何其他先前技术的变种，它也不是在其“前任”技术基础上稳步积累形成的。所以解释这种“新颖性”（novelty），即一种突变的、根本的新颖性，成了技术进化论者最主要的障碍。[5]根本性的新技术的出现，即相当于生物的新物种的出现，还不能被解释和说明。

一条很极端的出路，就是更刻苦地学习达尔文理论，然后声称，如果不同的设计者能够引发不同的技术“变异”，那么其中一些变异和它们背后的理念也许是根本性的。因此可以说，技术的进化可以是根本性的和突变性的，也可以是渐进式的。这听上去似乎很有道理，但是如果你

仔细观察在实践中如何实现根本性创新，就会发现这种说法根本站不住脚。雷达产生于半导体。你可以对20世纪30年代的半导体线圈进行任何你喜欢的改变，但是你永远不可能得到雷达。甚至你可以对半导体的理念进行你喜欢的任何改变，你依然得不到雷达。制造雷达需要和制造半导体不同的运作原理。

我不想对技术中的变异和选择置之不理。技术当然会存在于多个版本当中，出色的表现者也当然会被选择，所以后来的形式的确会按照先前这种形式繁衍下去。但是当我们面对主要问题，如根本性的新技术是如何产生的，这相当于达尔文理论中的生物新物种是如何产生的问题，我们就遇到了阻碍，达尔文机制就不好用了。

组合进化

有一种理解技术进化的途径，但是若要理解它，我们需要转换思维。我们真正要寻找的，不应该是达尔文机制如何对产生技术的根本新颖性起作用，而是“遗传”是如何作用于技术的。如果完整意义上的进化存在于技术中，那么所有的技术，包括新技术，一定是脱胎于之前存在的技术。也就是说，它们一定连接于、繁殖于某种之前的技术。换句话说，进化需要遗传机制——某种连接现在与过去的具体联系。从外部看（即视技术为黑箱的办法），是不可能看到这种机制的，就像我们很难说清激光是怎样脱胎于先前存在的技术一样。

如果我们从技术内部看技术会怎样呢？我们能看到任何能够告诉我们技术中的新颖性是如何作用的东西吗？我们能看到任何能产生一种适合技术的进化论的东西吗？

如果你打开一架喷气机引擎（专业名词叫作航空燃气涡轮发动机），你会发现里面的零部件，压缩机、涡轮增压机、点火系统。如果你打开在它出现之前的其他产品，你会发现同样的组件。20世纪早期的发电系统里面是涡轮和燃烧系统，同时期工业鼓风机单元内部是压缩机。技术继承之前技术的某个部分，所以把那些技术放在一起（组合起来），在很大程度上解释了技术是怎样产生的。这就使得根本性创新原本的不连续特征忽然显得没那么不连续了。技术在某种程度上一定是来自此前已有技术的新组合。

到目前为止，只有些许迹象能让我们用来解释新颖性。但是技术是如何被恰当地组建才是我讨论的核心。无论如何，新技术一定是产生于

已有技术的组合。

事实上，这个想法，就如进化论本身一样，也根本不新鲜。[6]它已经被不同的人讨论了不下100年。奥地利经济学家熊彼特就是其中之一。1910年，熊彼特27岁，他不仅直接关注组合和技术，而且关注经济中的组合。他说："生产意味着在我们的能力范围内组合材料和动力……生产别的东西，用不同的方法生产同样的东西意味着以不同的方式组合材料和动力。"经济中的变化产生于"生产方式的新组合"。用现代语言来讲，我们可以说它源于"技术的新组合"。

熊彼特之所以这么认为，是因为他一直在追问一个看似简单的问题：经济是如何发展的？我们也可以这样问：它是怎样进行结构性变化的？外部因素当然可以改变经济：如果发现了一种原材料的新来源，或者开始和一个新的外国伙伴进行贸易，或者开辟了新的领域，就能改变经济的结构。但是熊彼特问的是：经济能否在没有外界因素的情况下，完全从内部变革它自身，如果能，是怎样改变的？当代流行的学说，均衡经济学认为，那不可能发生。没有外界的扰动，经济将进入一个静止或者说均衡的状态，在均衡状态周围波动，并停留在那里。[7]但是，熊彼特认识到：有一股力量的源泉发端于经济内部，这种力量会打破任何经济内部的平衡。这种能量的来源是一种组合。经济体持续地通过组合旧技术来创造新技术，因此经济也不断地从内部扰动自身的结构。

熊彼特的著作直到1934年才被翻译成英文，而那时已经有其他一些学者在二三十年代得出了相同的结论：组合驱动变革——或者至少驱动技术创新。另外一位美国学派的成员，历史学家阿博特·裴森·厄舍（Abbott Payson Usher），在1929年时曾谈到，发明创造的进步，在于建设性地将先前的元素融入一个新的综合体中。[8]吉尔菲兰自己将其更简洁地描述为：一项发明是"此前技艺（art）的新组合"。在这种观念被抛出后，它偶尔被提起，但没有被援引，部分是因为没有人解释过这种组合是怎样带来新的发明的，熊彼特、厄舍、吉尔菲兰都没有，其他人也没有。这样说很简单：喷气机是弗兰克·惠特尔（Frank Whittle）和汉斯·冯·奥海因（Hans von Ohain）等发明者已有想法的组合，但是要解释这种组合是如何在原创者头脑里发生的，就不那么简单了。

组合至少为技术新颖性的诞生提供了一个思路。但是这仅仅是把具体某个新技术与之前存在的具体某个技术联系起来而已，但这并没有给我们呈现出全体技术从已有技术基础上构建起来的整体感。因此，我们需要加上第二层论述。如果新技术真是以前技术的组合，那么现存技术

的存量一定在某种程度上提供了新组合所需的部分。这样一来，以前技术的聚集就带来了进一步的聚集。

这种想法也有其自身的历史。一位与熊彼特年代相近的美国人威廉·菲尔丁·奥格本（William Fielding Ogburn），于1922年就提出过这种想法。[9]奥格本是一位社会学家，也是美国学派的成员。他当时着迷于是什么引起了社会的变迁（或者用他的话说，物质文明的变迁）这一问题。和熊彼特一样，他把由以前技术组合产生的发明当作变迁的源泉：似乎物质文化的装置越庞大，发明的数量越多。需要发明的东西越多，发明的数量就越多。这可以解释为什么更“原始”的社会没能发明出现代性的技术，他们没有必要的要素和利用这些要素的知识。有轨电车不可能从上一个冰河时期的物质文化中产生出来。蒸汽动力和机械技术时代使大量的发明成为可能。这里的洞见是了不起的。但是遗憾的是，它止步于此。奥格本没有用它来建构任何技术和技术进化的理论，而这对于他来说可能是易如反掌的。

> 技术思想前沿
>
> 组合进化：之前的技术形式被作为现在原创技术的组分，当代的新技术成为建构更新的技术的可能的组分。反过来，其中的部分技术将继续变成那些尚未实现的新技术的可能的构件。慢慢地，最初很简单的技术发展出越来越多的技术形式，而很复杂的技术往往用很简单的技术作为其组分。所有技术的集合自力更生地从无到有，从简单到复杂地成长起来了。我们可以说技术从自身创生了自身。这种机制便是组合进化。

如果我们将上述两部分放在一起，即新技术产生于已有技术的组合，以及（因之而来的）现存技术会带来未来技术，我们是否能够得到技术进化的机制呢？我的回答是肯定的。该进化机制可以简述如下：之前的技术形式会被作为现在原创技术的组成部分。当代的新技术将成为建构更新的技术的可能的组分（构件）。反过来，其中的部分技术将继续变成那些尚未实现的新技术的可能的构件。以这种方式，慢慢地，最初很简单的技术发展出越来越多的技术形式，而很复杂的技术往往用很简单的技术作为其组分。所有技术作为一个整体自力更生地积少成多、由简入繁地成长起来了。我们可以说技术从自身创生了自身。

我将这种机制称为依靠组合而形成的进化，或简称为组合进化

（combinatorial evolution）。

当然，仅仅如我以上所述还不完全。组合不可能是技术进化背后唯一的机制。如果它是的话，现代技术，比如雷达或者核磁共振成像（即医院用的MRI）应该可以从弓钻和制陶技术，或者任何我们认为存在于技术发端之时的技术直接创生而来。这样我们就要面对一个问题，即确定这个发端的时间。如果弓钻和制陶技术本身也是由更早的技术组合而来的，那么这些元初的技术（ur-technology）是从何而来的呢？由此我们将陷入无穷的回溯之中。一定有超出组合之外的其它事物在持续创造新的技术。

我认为，这些超出组合之外的其它事物，就是持续地捕捉新的自然现象，并为了特定目的而利用这些现象。在雷达和MRI的例子中，所利用的现象就是电磁波和核磁共振的反射，其特定目的分别是探测飞机和诊断疾病时对人体进行成像。技术的建构不仅来自已有技术的组合，还来自于对自然现象的捕捉和利用。在技术时代发端之初，我们只是直接地识别并利用自然现象：火的灼热、片状黑曜石的尖利、运动中的石头的冲力……我们所有的收获都来自于对这些现象的掌握以及对它们的组合。

像这样简单地进行描述比较容易，但是进一步准确、详细地描述则需要更多的工作。我将不得不加以说明，“新技术是之前技术的组合”到底是什么意思。技术不是随机地对现存技术进行组合而成的。所以我需要提供详细的原理，来解释组合是怎样工作的。我们要解释像涡轮喷气发动机这样的技术是如何作为现存技术的组合而产生的。再退一步，这表明，我们需要研究技术是怎样被逻辑地结构化的，因为组合，无论它是如何发生的，肯定都要依据那种结构而发生。我们必须要关注人类，特别是人类思维在这一组合过程中的巨大作用。新技术先是精神的建构，之后才是物质的建构。这一精神过程需要仔细探究。我们必须关注技术到底是如何变成了现实：人类的需求是怎样召唤出新技术的创造。我们必须弄清楚技术怎样创造出技术，即新技术从已有技术整体中涌现出来。为了回到问题的根本，我们必须清晰地定义所谓的“技术”。

本书的主题

本书主要论述的是：技术是什么，以及它是如何进化的。我们试图建立一门关于技术的理论，也即可以用来解释技术行为的“一组自洽的一

般命题”。[10]本书尤其想创建的是关于技术的进化理论。

我计划从完全空白的状态开始，将技术的所有相关项都不视为理所当然的。我将基于三条基本原理逐步建构这一理论。第一条我已经谈到过：技术（所有的技术）都是某种组合。这意味着任何具体技术都是由当下的部件、集成件或系统组件建构或组合而成的。其次，技术的每个组件自身也是缩微的技术。这听起来很奇怪，我将会对这种说法进行辩护。但现在只需要把这句话理解为，由于组件本身和整套技术一样，也是为某个特殊目的服务的，这些组件本身也是技术。第三条基本原理是，所有的技术都会利用或开发某种（通常是几种）效应或现象。

我会在以后的篇幅展开讲这几项原理。但注意，这三条原理立刻会带给我们一种内部的视角来看待技术。如果技术是组合而来的，那么它们就有了一幅内在景象：技术组件的集成或组合是为了满足它们的目的。这种内在性由本身也是技术的组件或子系统构成的。我们可以发现，新颖的技术是借由组合已有的技术来触发的，当然这也需要捕捉现象；我们可以发现，通过改变技术内部的组件，用可以改善性能的组件来替代，从而推动技术进步；我们可以把不同的技术看做是用同样的组件构成的，这些组件同样是从之前的技术传承而来。从这个视角开始观察技术，会发现一种技术的“遗传学”。当然这不等价于DNA或细胞，也没有那样美丽的秩序，但是它依然呈现了一种丰富的相互联系的世系。

所有这些听起来很有机——非常有机，其实这一视角也会把我们带向既是机械性的、也是生物性的思路上。可以肯定的是，技术不是生物有机体；几乎从定义上就可以确定，技术是机械性的：不论是排序算法还是原子钟，它们的内部组件都是以可预期的方式来互动的。但是一旦我们将技术以不断组合成新组合的方式展现出来，我们就不太能将其看作仅仅是像发条装置那样的独立部件，而是各种工序构成的复合体，与其他复合体交互构成新技术的复合体。这样我们就发现了一个美丽新世界，在这里，作为一个整体的技术从已有技术中形成了各种新技术。技术有机地从内部建造了自己，而这将是本书的主题之一。

这里有个视角的变化，要从将技术看作是有固定目的的独立对象，转变为将技术看作是可以无限构成新组合的对象。这种视角的转换不仅仅是抽象的，它实际上是技术角色目前正在经历的一个大转换的真实反映。那种标志着制造业经济的旧有的、工业的过程技术，如平炉炼钢过程、提炼原油的裂解过程，确实在很大程度上被固定了。它们只在一个固定的地点生产一种东西，它们加工某种特定的原材料、产出特定的工

业品，并且大部分是在分开的、独立的工厂中进行的。但是现在这些相对独立的生产技术已经开始让位给不同形式的技术。这些技术可以很容易地被组合，形成技术模块，可以被反复地使用。全球定位技术可以直接提供方位，但是它几乎不能独立工作。它被当作一个要素，去和其他的要素组合共同为飞机和轮船导航、辅助土地勘探、管理农业生产。它就像一种化学中高度活性的成分（例如，氢氧根离子），自己虽然只参与一点，但却主导着不同的组合。这种说法也适用于数字革命中的其他元素：算法、交换机、路由器、中继站、网络服务等。同样地，我们还可以如此这般地描述组成现代基因工程或者纳米技术的元素，它们可以在无限的组合中匹配在一起，可以因为不同的目的被装配、再装配，它们也为持续的组合提供可用的基础构件。

现代技术不仅是稍具独立的生产方式的集合，而且已经进化成创造经济结构与功能的开放语言。以十年为计算单位来看，慢慢地，我们从生产固定的物理产品的技术转变成，为了新的目的可以进行无限地组合和装配的技术。

技术，曾经的生产手段，正在成为一种“化学物质”。

在试图总结出一种技术理论的过程中，我们面对的第一个挑战是去看看我们是否能够从一般意义上谈论技术，这是我们不可逾越的一个前提。我们可以随机选3种技术：水利发电、塑料铸模工艺和养蜂，它们似乎无任何相似之处，但是我们将会在下一章看到，在这些技术结合成一体的方式上，实际上存在着同样的逻辑。我们会看到这种组合是如何运作的、技术是如何形成的、技术又是如何发展的，以及技术将如何进化。

但是首先我们必须解决一个更根本的问题，技术到底是什么呢？

THE NATURE OF TECHNOLOGY

第2章

组合与结构

新技术都是在现有技术的基础上发展起来的，现有技术又来源于先前的技术。将技术进行功能性分组，可以大大简化设计过程，这是技术“模块化”的首要原因。技术的“组合”和“递归”特征，将彻底改变我们对技术本质的认识。

当我们说“技术”的时候，我们在说什么？[1]令人烦恼的是，无论是从字典中，还是从技术思想家的著述里，我们得到的答案都很模糊。他们说，技术是知识的分支，是科学的应用，是一门对技术的研究，是一种实践，或者是一种活动。牛津词典以一种可爱而又一本正经的方式解释道：“技术是机械艺术的集合，它是一种文化用来运转经济和社会的资源。”可以推测，这里所说的“机械艺术”是指特定的方法、实践或者装置，它们被某种文化加以利用，从而使其发挥经济或社会功能。

所有这些说法可能都没错，况且词语本身通常就有多重含义。但是如果我们接受了以上说法，技术难道就真的是“知识”[2]，或者是“科学的应用”，或者是一门学问，或是某种“实践”、某个“集合”吗？能同时涵盖所有这些含义吗？概念界定之所以事关重大，是因为，我们如何看待技术将决定我们如何看待技术的产生与形成。如果技术是“知识”，那么它一定和知识有同样的起源方式；如果它是“实践”，那么它一定是通过“实践”产生的；如果它是“应用科学”，那么它一定以某种方式从科学中衍生而来。如果这些定义隐含着我们对技术的理解，那么事实上它们并没有被很好地融合起来，甚至是相互矛盾的。

为此，我们需要深入其中。我将回到第1章所说的三个原理，并着手从头开始定义技术。

技术思想前沿

技术的三个定义：

- 技术是实现人的目的的一种手段。
- 技术是实践和元器件的集成。
- 技术是可供某种文化中利用的装置和工程实践的集合。

首先，我将给出本书将会用到的三个技术定义。

第一个也是最基础的一个定义：技术是实现人的目的的一种手段。对于某些技术（例如，炼油）来说，其目的是不言而喻的；而对于另一些技术（例如，计算机）来说，其目的可能就比较模糊又多重，甚至是在不断变化。作为手段，一项技术可能是一种方法、过程或者装置，比如一个特定的语音识别算法，或者化学工程中的过滤法，或者柴油发动

机。技术也可以是简单的，比如一个滚动轴承；也可能是复杂的，比如波分多路复用器；它们可能是物质性的，比如发电机；也可能是非物质性的，比如一种数字压缩算法。但无论如何，技术总是完成人类目的的一种手段。

我会用到的第二个技术定义是复数性质的：技术是实践和元器件的集成（assemblage）。比如电子技术和生物技术等，它们是许多技术和实践构成的集合或者工具箱。严格说来，我们应称它们为技术体（bodies of technology）。但是由于用“技术”来指称“技术体”已经非常普遍，因此我将继续采用“技术”这一称谓。

我用到的第三个定义是：将技术视作可供某种文化中利用的装置和工程实践的集合。牛津词典称之为“机械艺术的集合”，韦氏词典表述为“人类创造物质文化的手段的总和”。当我们指责“技术”加速了我们的生活，或者说“技术”是人类的希望时，我们指的就是总体意义上的技术。有时这种含义逐渐变成了一种集体活动，比如，当我们说“技术就是与硅谷相关的一切”的时候。我把这也当作整体意义上的技术的另一种说法。技术思想家凯文·凯利（Kevin Kelly）称这个整体为“技术元素”（technium）。我喜欢这个词。但是在本书中我还是使用惯常的说法——“技术”。

我们之所以需要以上三个定义，是因为它们的意义不同，所属的范畴不同。每个范畴下的技术的形成和进化都不同。作为一项单数意义上的技术（technology-singular）——例如蒸汽机，是作为一个新的概念而产生的，并通过修正它的内部构件得以发展；作为一项复数意义上的技术（technology-plural）——例如电子，则往往通过围绕某些现象和器件建构起来，并通过改变它的构件和实践而得以发展；而作为一般意义的技术（technology-general）——所有过去和现存技术的总和，则产生于对自然现象的应用，并随着由旧要素组合而成的新要素的形成，有机地成长起来。

在本书中，我将会更多地涉及第二个和第三个技术范畴，特别是关于技术的整体是怎样进化的问题。但是，本章我将集中讨论单数意义上的技术，因为技术集合是由单数的技术，即一个一个的技术组成的，因此我们需要确切地弄清楚它们到底是什么，以及什么是这些单数意义上的技术所共享的逻辑。

技术思想前沿

- 技术是实现目的的一种手段，它是一种装置、一种方法或一个流程。
- 技术提供功能，功能指技术要执行的某一类任务。

如前所述，技术是完成目的的一种手段：它可能是一种装置、一种方法或者一个流程。一项技术就是要做点什么，要执行一个目的。为了强调这一点，我有时会把技术看作是可执行的（executable）。但这在我们开始进行实际的阐述时，会有一点麻烦。比如把铆接机看成是可执行的技术很容易：它被“激活”就是要完成一项特定任务。但是我们如何对待那些好像不能被“激活”的技术呢？比如一座桥——一座桥是可执行一个目标的吗？那么一座大坝呢？我的回答是，这里的每一项技术都有一个进行中的或者预设的任务要执行。桥要承载交通，大坝要储水或提供能量。如果不出意外，这些技术都应该时刻发挥效力。在这个意义上，每个技术都在执行任务，所以说技术是可执行的。

还有一个贯穿全书的词是实用功能（functionality），技术提供实用功能。实用功能是指技术要执行的某一类任务。例如GPS（一种全球定位系统），定位就是它的功能。GPS有许多特殊目的：飞机导航、地面定位及巡航，但是当我们在一般意义上讲GPS的目的时，GPS就是提供定位功能的一种手段。

这样的解读好像已经可以了，但是关于技术的定义还是显得有些混乱。实现目的的手段可能是装置、方法或者流程，它们看起来完全不同，以至于会令人觉得我们所谈论的根本不是一回事情。真的是这样吗？我们试一下，先将方法和流程归为一类，因为它们都是指通过一系列阶段、步骤去转换某些东西，因此具有逻辑上的相似性。但是装置和流程，比如半导体收音机和石油冶炼流程，看起来就完全不同了，一件装置似乎就只是一个硬件，而完全不像是一个流程。但是这仅仅是表面上的认识，一个装置总是按照某个流程处理某件事情，它按过程从头至尾完成既定任务：飞机是“按流程处理”（processes）旅客或者货物，完成将他（它）们从一地运往另一地的任务。如果我们想扩展这个想法的话，一把锤子则可以表达为“按流程处理”一个钉子。

半导体收音机也是“按流程处理”的。[3]它接收无线电信号，并通过输入天线将其转换成微电压差；接着通过一个共振装置，从信号中（例如，某一个特定的电台）选取一个特定的频率；再将结果通过一系列放大步骤进行传递；然后分离出声音或音乐（原声的）信息，将上述结果

再一次放大，并把它传入到扩音器或者耳机中。因此，可以说半导体“按流程处理”信号。在这里，我是完全按照半导体工作原理实际发生的过程进行描述的，而不是隐喻式的。它从空气中抓取信号、提纯它们，将它们转换成声音。这是一个微缩的提取过程，一个我们可以放在掌心的微型工厂。所有的装置实际上都在“按流程处理”着什么。这就是为什么经济学家始终认为技术就是生产手段的原因。

那么反过来，我们能将方法和流程看作是装置吗？回答是肯定的。流程和方法，比如炼油或者分类算法，都是操作的序列（sequences of operations）。但是为了执行，它们常常需要一些硬件、一些物理设备才能真正完成。我们可以把这种物理设备看成是正在执行一系列操作的“装置”。当然，在炼油的例子中，这个装置显然非常巨大。如果我们将执行流程的各种装备也包括进来的话，那么流程就是装置。这样看，炼油和半导体收音机就没什么两样了。两者都是“按流程处理”，只是一个是用大型工程设备，另一个是用小型的电子元件而已。

还有一种更一般的方式可以用来表明装置和流程的确可以归于相同的范畴。技术包含一系列操作，我们可以称之为技术的“软件”。这些操作需要物理设备去执行，我们可以称之为技术的“硬件”。当我们强调“软件”时，我们看到的是过程和方法；当我们强调“硬件”时，我们看到的是物理设备。实际上这两个方面都属于技术，但如果只强调一个方面，而忽视另一面，则会使它们看起来好像分属两个范畴，而这只是从不同侧面看待技术的结果。

最后一个想法，实际上是一个问题。当我们谈论“技术”时，比如斜拉桥（例如，法国的诺曼底大桥），我们是将它看作装置还是方法？抑或是在谈论那个装置或方法的“理念”（例如，斜拉桥的概念）？如果我们意识到，在我们对事物进行抽象和表征的时候，这种情况经常发生，我们是可以给出答案的。比如，当我们谈论一只黑腹缕的时候，有时我们是真的在说一只鸟——一只沿着海岸线欢跑的鸟；有时我们谈论的是一种鸟的种属，一种鸟类物种的概念（鸻科）。在这种情况下，依据不同的语境，在实体和概念间转换是完全正常的。

我们同样可以这样来分析技术。我们可以谈论德国克雷菲尔德市附近一个制氨厂所用的哈伯制氨法过程，也可以只将哈伯制氨法作为技术的理念或概念来谈论，而且我们还可以按照需要在现实和抽象之间来回转换。连带的好处是，如果我们接受了把某种事物当作抽象的概念，我们就很容易在概念层面上对技术进行拉近和推远景深的探究。我们可以

谈论“波音787客机”这种特殊的技术，也可以谈论“客机”这个一般意义上的技术，还可以谈论“飞行器”这个更加一般意义上的技术。当然，严格意义上“飞行器”是不存在的。但是我们发现抽象概念很有用，比如由于“飞行器”这个概念表明了某些共同部件和构架，因此在评价和谈论这类技术时就可以使用这个词。[4]

技术结构的形成

现在我想回到我在第1章末尾抛出的那个问题，各种技术之间的逻辑是共通的吗？它们在组成方式上有相同的结构吗？

如我所说，答案看起来并不乐观。虽然如此，我们还是能看到技术确实分享了一种解剖学意义上的结构。技术在这方面就像某类动物——比如，脊椎动物。脊椎动物无论在结构样式还是外观上都相当不同。非洲河马和一条蛇看起来毫无相像之处，但是它们都有分节、相互连接的柱状结构，以及心脏、肝脏、肾脏和神经系统；它们都左右对称，且全都建立在细胞基础之上。技术的共同结构当然不在于拥有共同的器官。但是无论共同拥有的是什么，它都将是我们讨论的核心，因为它涉及技术怎样被放到一起，是怎样形成的。就像俳句(5)，其结构过程一定就是指音节是如何被放到一起并最终呈现出俳句形式的。

那么，这些结构在解剖学意义上有什么共同之处呢？

一开始，我们可以说技术是由各组成部分或元器件放置或组合（联结）在一起的。因此技术是为达到某种目的的一种组合。这一组合的原理是我在第1章提到的三个原理之一。一台水力发电机集中了几种主要的构件：一个用于储水的蓄水池，一个带有控制阀门和被称为水渠的进水管组成的进水系统，由高位能水流驱动的涡轮机，由涡轮机驱动的发电机，将输出能量转换为高压电的变压器以及排水用的泄流系统或尾水渠。类似的构件或者子系统、子技术（或对于过程技术来说的阶段）是一组能够独立的，可以与其他构件相分离的构件。当然，一些非常基础的技术，例如一个铆钉，可能只有一个部件。我们也可以认为，如此基础的技术也是一种“组合”，就如同数学家允许一个集合里只有一个数一样。再进一步，我们甚至可以承认拥有0个部件且没有目的的技术。为了配合正确的数学表达，就像数学里存在空集一样，我们应该承认拥有0个部件且没有目的的技术。毫无疑问早就有人这样想过了。

结构首先是指技术是由零部件构成的。如果我们观察到技术总是围

绕着一个中心概念或原理，如“事物的方法”或者一个可行的核心理念来组建的话，我们就能看到更多的结构。钟的原理就是依据某个固定频率的节拍来计数。雷达的原理（核心理念）就是发送高频无线电波，然后通过分析来自物体表面的反射信号来探察远处的物体。激光打印机的原理是用计算机控制的激光在硒鼓上“写”映像，再应用基本的电子照相法，使墨粉微粒以静电书写方式粘到硒鼓上，之后再熔融到纸上。

> 技术思想前沿
>
> 技术的最基本结构，包含一个用来执行基本功能的主集成和一套支持这一集成的次集成。

为了能够成为物理实在，一个原理需要以物理部件的形式被表达出来。在实践中，这就意味着一项技术包括一个主要组合体，即一个装置或者方法的基础或支柱，用这个组合体来执行基本的原理。这个支柱由另外一些组合体来支持，这些组合体被用来支持主要组合体的运转，调节它的功能、提供动力以及执行其他次一级的任务。所以，技术的最基本结构，包含一个用来执行基本功能的主集成和一套支持这一集成的次集成。

让我们用这种方式来看一下喷气式发动机。它的原理很简单：在稳定的压缩空气流中燃烧燃料，产生高速气体向后排出。根据牛顿第三定律，这会产生一个等效反作用力。为完成任务，机器会采用一个主集成，包括五个主要的系统，进气道、压气机、燃烧室、涡轮和尾喷管。空气由进气道进入，流进由一系列大风扇组成的压气机加压，高压气流进入到燃烧室中，和燃料混合并被点燃，之后高温气体驱动涡轮运转，涡轮驱动增压机，高温气体从尾喷管高速喷出，进而产生推力。（现代的涡轮风扇发动机前面有一个巨大的风扇，也是由涡轮增压机驱动的，也能产生很大的推力。）

这些元器件形成了核心集成，围绕在它周围的是支持这一主功能的许多非常复杂的子系统，主要包括：燃料输送系统、压缩机反熄火系统、涡轮叶片冷却系统、发动机仪表系统、电气系统等。所有这些集成件和子系统互相配合：燃料供应系统的输出变成了燃烧系统的输入，并通过仪表系统把压气机的表现表征出来。为了支撑这些功能，集成件之间通过复杂得像迷宫似的管道和电线相互连接来传递功能，以便让其他的系统得以应用。

这一建构过程和计算机编程没什么不同。都是要应用一个基本原理，即核心概念或者逻辑，作为编程背后的支撑。完成这个原理需要一套具有积木一样结构的支撑程序和子程序系统。这在计算机中，常恰如其分地被称为“主程序”。接下来，“主程序”又需要其子功能或子程序的支持。要在计算机屏幕上建立一个图文视窗程序，需要一系列的功能去完成，如规定尺寸、确定位置、展示题头、发表内容、将其放在其他窗口的前面、删除完成的任务等。子功能之间互相需要，以使彼此相互契合。各个子程序就处于不断的相互联系之中，它们不停地“对话”，就如同喷气发动机那样。

实际看来，一台喷气式发动机和一台计算机是很不同的东西。一个是一套物质零件，另一个是一套逻辑指令。但是它们的结构却是相同的：都是由集成块结构起来，集成块之间相互联系，共同服务于一个执行某一个基本原理的核心集成，再辅之以其他的互动的集成子系统或组件系统的支撑。

不论是在喷气式发动机还是在计算机程序里，所有的组件之间必须小心地保持平衡。每个部分都必须能够在一个由其他相关部分设定的约束范围内运行，包括温度、流速、荷载、电压、数据类型以及协议等。并且反过来，每部分又都要为依赖于它们的那些组件建立一个适合的工作环境。这意味着在实践中进行权衡一定非常困难。

每个模块或者内容一定要提供刚好正确的动力、尺寸、强度、重量、措施或数据结构来适应其他模块。因而，每个部分都要被设计得与其他部分能够平衡地匹配。

这些不同的模块和它们之间的联系共同形成了一个工作构架（working architecture）。理解技术意味着理解它的原理，以及它是如何将这种原理转化到工作构架中去的。

为什么要模块化

我们现在已经为技术们建立了一个共同的结构，即它们是由零部件组成的组件系统或模块。其中一部分形成了核心集合，其他部分行使支持功能，它们自己可能还有子集合和次级零部件。当然完全没有任何法则规定技术的组分一定要被聚集成集合或功能性的集团。例如，我们可以很容易想象一种完全由单一元器件放置在一起组合而成的技术，然而除了在极端情况下，几乎是没有那样的技术的。

为什么会这样呢？为什么技术应该是从集成件或者单个零部件中被组织结构起来的呢？

> 技术思想前沿
>
> 将技术的构件模块化可以更好地预防不可预知的变动，同时还简化了设计过程。但只有当模块被反复使用且使用的次数足够多时，才值得付出代价将技术分割为功能单元。

几年前，赫伯特·西蒙（Herbert Simon）[6]讲了一个关于两个制表匠的经典寓言故事。假设每只表都集成了1 000个零件。一个名叫坦帕斯的钟表匠一个零件一个零件地安装，但是，如果他的工作被打断了，或者丢下一只没完成的手表，他就必须从头开始。相反，另一个名叫赫拉的钟表匠则是将10个模块组装在一起，最后组成手表。每个模块又由10个子模块组成，每个子模块再由10个零件组成。如果他暂停工作或者被打断工作，他只是损失了一小部分工作成果。西蒙的重点是：将零件集成化可以更好地预防不可预知的变动，且更容易修复。对此，我们可以进一步加以扩展，模块将允许技术的组成部分分别进步：可以对每个部分分别加以关注和改进，对工作性能分别进行试验和分析——每个“集成”可以“悄悄地”被探察或者更换而不必解体余下的技术整体。而且这样做还可以允许通过技术的重新配置来适应不同的目的，不同的组装可以根据需要被来回变换。

将技术进行功能性分组（functional groupings）还简化了设计过程。如果设计者要面对数以万计的零件，那么他们将被淹没在细琐零件的汪洋之中。但是如果能够将技术分割成不同的建构模块（例如，计算机的计算程序、记忆系统、电源系统），设计者就比较容易加以记忆并分别给予关注，从而也较容易看清这些大一些的零件如何能够互相匹配、共同服务于整体。将技术分割成组或模块，有点像认知心理学中的“组块”（chunking）概念。[5]我们用它来将复杂的事物（例如，第二次世界大战）分解成更高阶的部分或组块（战争的导火索、战争的爆发、苏联的入侵、太平洋战争等），这样我们就能更容易理解和运用它们了。

将技术分割为功能单元需要付出一些代价，至少要有一些精神上的努力。只有当模块被反复使用，且反复使用的次数足够多时，才值得付出代价将技术进行分割。这和亚当·斯密的劳动分工理论类似：斯密指

出，只有在生产的数量足够大的情况下，才值得将工厂的工作划分成专业工作。[6]我们可以说，模块化（modularity）之于技术经济（technological economy），就如同劳动分工之于制造业经济一样。一项技术被应用得越多，分成的模块就增加，经济也因之而发展。或者可以用斯密的话来表达同样的意思：技术的分块随着市场的扩展而增加。

随着功能单元被更多地应用，其被组织的方式也会发生变化。一个模块或集成开始成为一种单个零件组成的典型的松散集团，可以联合执行一些功能。后来，这个集团固化成一种特殊的结构单元。例如，DNA扩增（DNA amplification，一个将小的DNA样本复制成亿万样本的过程）在最初只是一种实验室技术的松散的组合，目前它已经内嵌于一种专用结构机体内部了。这里有一条通用的规则：一开始的一系列松散地串在一起的零件如果被用得足够多，就会“凝固”成独立的单元。技术模块随着时间的推移会变成标准组件。[7]

递归性及其重要性

我们现在看到很多技术内部的结构了，但事情还远远没有结束。结构还有另外一方面。根据我们的组合原理，技术包含如下组成部分：集成体（assemblies）、系统、单一零件（individual parts，即不可再分的部分）。因此我们可以在概念上将技术从上到下分解为不同的功能组件（functional components，忽略它们是支撑性的还是核心性的），从而将技术分解成主集成（main assemblies）、次级集成（subassemblies）、次次级集成（sub-subassemblies）等，直至分解为最基本的部分，这样会让我们整体地了解技术。

> 技术思想前沿
>
> - 技术具有层级结构：整体的技术是树干，主集成是枝干，次级集成是枝条，最基本的零件是更小的分枝。
> - 技术具有递归性：结构中包含某种程度的自相似组件，也就是说，技术是由不同等级的技术建构而成的。

以这种方式形成的等级呈树形结构[8]：整体的技术是树干，主集成就是枝干，次级集成是枝条等，基本的零件是更小的分枝。当然这不是一株完美的树：枝干和枝条（主集成和次级集成）会在不同的层次交叉

勾连、互相作用。树形结构的层级数取决于主干上的枝条，以及那些有代表性的小分枝的数目。对于坦帕斯来说，技术可以分为2级：整个手表和单个零件；而对于赫拉来说，技术可以分为4级：手表、主模块、次级模块和单个零件。真实世界的技术可以有任何数量的层级数：从2级到10级或者更多。技术越复杂或越模块化，层级就越多。

到目前为止，除了层级，我们还没有关于结构的更一般的概念。但是，我们可以再多说一点儿。每一个集成或次集成都有一个要执行的任务，如果没有那个任务，它就不会存在。因此，每个部分都是一个目的的手段。因此，就如我先前的定义，每个部分都是一个技术。这意味着，每个集成、次级集成和单个零件都是可执行的——都是技术。于是就出现了这样的结果：技术包含的集成块是技术，集成块所包含的次一级的集成块也是技术，次一级集成块包含的再次一级的集成块还是技术。这样的模式不停地重复，直到最基础水平的基本零件为止。换句话说，技术有一个递归性结构，技术包含着技术，直到最基础的水平。

递归性将是我们要讨论的第二个原理。[9]在数学、物理学或者计算机科学之外，“递归”并不是一个广为熟知的概念，其本意是指，结构包含某种程度的自相似组件。当然，在我们的语境中，它并不意味着一架喷气式发动机内部包含着一些小的喷气式发动机，那将是荒谬的。它只是简单地意味着一个喷气式发动机（或者更一般地讲，任何技术）包含的构件也是技术，并且这些构件包含的次一级构件也是技术，并以这样的方式不断重复或再现。

如此说来，技术是由不同等级的技术建构而成的。这暗示我们应该怎样思考技术，也意味着无论我们在怎样的一般意义上谈论单个技术，其实也包含着更低层次的模块或者次级系统。特别是，如果一项技术包括主要模块和支撑模块，那么每个模块或子系统也一定是按照这种方式被组织起来的。

到目前为止，递归性听起来还是比较抽象的。但是当我们在行动中审视它的话，它就会完全变得具体了。我们可以考虑一些复杂的技术，例如F-35“闪电”II型战斗机。（军事案例很有帮助，是因为它们能提供更多的层级。）

F-35是一种常规战斗机，短道/垂直起降，并且可以舰载起飞。我脑海里出现的是舰载的那种战斗机——F-35C。它是单引擎、超音速隐形战斗机。这意味着它吸收的雷达信号非常少，因此难以被发现。F-35C拥有倾角表面，具备隐形飞机的典型特征。然而整体上看它的造

型还是很优美的，拥有巨大的翅膀，其大部分与机身和尾部相连，机尾是从后面突出来的两个部分，上面固定着互相垂直成V形的稳定器。

设计这些细节的原因是，F-35C在设计上需要实现一系列互相冲突的目标。它需要在结构上足够坚固并且足够重，这样才能承受发射载体和尾钩降落的高强度；同时它还需要保持高机动性和远程燃料供给性能；它需要有优异的低速控制能力以完成在航母上降落，又要能以超过音速1.6倍的速度飞行；它需要有多棱折面的外表以躲避雷达的追踪，同时要能够正常飞行。

F-35C型战斗机有着多重目的：提供近距离空中支援、拦截敌机、压制敌方雷达防御、消灭地面目标。因此，它是一种手段，一项技术——具有可执行性。

如果我们继续追随F-35C型战斗机的层级树到达它的末梢，我们会看到什么？我们可以将它分解如下：机翼与机尾、动力装置（或者发动机）、航空电子控制系统（或者飞机电子系统）、起落装置、飞行控制系统、液压系统等。如果我们单独观察动力系统，我们可以将其分解成常规喷气发动机的子系统：进气装置、压缩机系统、燃烧系统、涡轮系统、喷嘴系统。如果我们只看进气系统，则可以看到，在机身两侧、机翼前面的位置，有两个箱式超声速进气道。超声速进气道调整着进气速度，这样发动机就能在起飞速度和飞行速度为1.6马赫时保持同样的进气速度。为此，标准超音速进气道会内配一个移动盖板。而F-35C则会更聪明地设计一个凸起，并称之为DSI（无分流超音速进气道）-凸起，位置在机身靠近进气口的地方。这个凸起可以预调气流并有助于控制冲击波。我们还可以进一步深入下去：DSI-凸起这个组件又包括某些合金次结构。

我们已经到达了等级的最底层，最基础的可执行水平。请注意在“单个构件”这个术语背后的循环主题：包含着系统的系统之间相互勾连、相互影响、相互制衡，包含着可执行性的可执行性，以及包含着技术的技术。层级式的安排贯彻始终，直到单个元素为止。

我们也可以向上追溯这些可执行等级。F-35C是一个更大的系统（舰载飞行联队）的一个执行部分。舰载飞行联队又包括几个歼击机中队以及其他的后勤飞机，其目的是提供攻击力和电子战争能力。舰载飞行联队又是更大系统的一部分：它被部署在航母上。航母也是一个执行部分：它的目的是收留、发送并接收飞机。同时，航母也是更大系统（航空母舰战斗群）的一个执行部分。顾名思义，这是一个围绕在航空

母舰周围的船的组合：导弹巡洋舰和护卫舰、驱逐舰、补给船和核潜艇。它们各自都有自己的目的（军事术语称作“任务”）：指明目标、发动攻击、提供给养，因此航母舰队也是可执行的。对于航母舰队来说，它也可能是更大系统的一个部分，比如说集团战区：由航母舰队和陆基航空单位、空中加油机、海军卫星侦察、地面监控，以及海洋航空单位的支持共同构成。这一更大规模的集团战区可能是流动变化的，但是它依然是一个工具——它执行海军作业。所以它也是可执行的。如果我们愿意以这种方式看待它，那么它也是一个技术。

因此我所描述的整个战区系统是一个可执行的“技术”层级结构。它由9层甚至更多的层级组成。我们可以从任何一层（任何一个组件系统）进入这个技术，并且发现它也是一个运转着的可执行的层级结构。这个系统是自相似的。当然不同层次的可执行体（executables）在类型、主题、外观和目的上是不同的。空气进口系统绝对不会是F-35C战斗机的微型版本。但是在任何层级，每一个系统以及子系统都是一个技术、一种手段以及一个可执行文件。这样看来，自相似性上下传递的是一个拥有许多层的递归性层级结构。

我们可以从这个例子中得出许多一般性的结论。

技术思想前沿

在真实世界中，技术是高度可重构的，它们是流动的东西，永远不会静止，永远不会完结，永远不会完美。

所谓标准的观点，我指的是大多数技术思想家持有的观点，认为技术在很大程度上是自给自足的，并且在结构上是固定的，偶尔会有一些创新。但这种观点只有我们在封闭的实验室里进行抽象思考时才成立。“在野外”（我的意思是在真实的世界里），一个技术绝少是固定不变的。它会不断地变换结构，当目的改变时，它会去适应并进行重新的配置，之后改进就发生了。一架舰载喷气式战斗机前一天可能充当着一个独立的组件，后一天可能就被指定去保护雷达预警机了，从而成为一个新的临时分组的一部分。如果需要，新结构、新架构可以在任何层级迅速且容易地形成。在真实世界中，技术是高度可重构的，它们是流动的东西，永远不会静止，永远不会完结，永远不会完美。

我们也倾向于认为，在经济的世界中，技术是在一定规模层面上存在的。如果是传统技术过程（纯氧炼钢法），那主要就是在工厂规模上

的技术；如果是一个设备（移动电话），就主要在产品规模上。但是在我们的例子当中，集团战区整体可以作为一个技术，同样，在它的飞机控制系统中的最小的一个晶体管也是一个技术，而这两者之间所有的元器件也都是技术。因此我认为，对技术来说，没有什么特征性的规模。

这并不是说在技术中没有“高”、“低”之分。处于高层级的技术指导或者通过“编程”（正如计算机中的程序）对处于低层级的技术下达指令，并组织它们执行任务。航母按程序指令，让它的飞机去执行它的任务。低层级技术决定高层级技术可以完成什么。集团战区受限于它的航母的能力极限，航母则受限于它的飞机的极限。

我们也可以看到，每个技术都至少是潜在地准备好去成为高一层级技术的一个零部件。一架F-35C在原理上可以作为单机完成单一任务的技术，但是它同时也可以作为一个更大的系统舰载空军联队中的一个组成部分，它在这个背景下执行它的任务。这有力地支持了我们的主张：所有的技术可以作为组件为其他的新技术作好准备。

递归性还有一个更深的含义。在技术世界，一个非常普遍的现象是，某一层级的变化一定要与其他层级的变化相协调。F-35C提供了一套不同于它的前任F/A-18大黄蜂的能力。而这意味着，处于比它们高层级的控制和部署它们的航母系统也必须改变自身的组织方式。在以后的章节中，我们将涉及更多关于技术需要在不同的层级中变换的主题，因为这种变换是存在于所有技术当中的。

技术思想前沿

组合不仅仅是将具有匹配的概念或原理的目的聚集起来，它还需要提供一套主要的集成件或模块去执行那个核心理念。为此它必须提供进一步的集成，进而需要更进一步的集成来支撑。而所有的这些零部件和集成件必须合在一起才能奏出和谐的乐章。组合必须是高秩序性的过程。

我在本章讨论了技术的逻辑结构，但是我不想夸大技术间的共同之处。一个航母舰队是一个技术，但是它和威士忌蒸馏过程是非常不同的。尽管如此，我们还是同意独特的技术特征分享共同的解剖组织（anatomical organization）这一观点。每一种装置都来自于一个中心原理，并且有一个中央集成（一个装置的整体骨干或者执行方法），加上其他的零部件围绕其周围，令其可以工作并且规定它的功能。这些组

件的每个自身都是一个技术，因此它自己也有一个核心骨干，以及附着其上的其他组件。这一结构是递归性的。这种结构的存在带给我们一个重要的启示：组合不仅仅是将具有匹配的概念或原理的目的聚集起来，它还需要提供一套主要的集成件或模块去执行那个核心理念。为此它必须提供进一步的集成，进而需要更进一步的集成来支撑。而所有的这些零部件和集成件必须合在一起才能奏出和谐的乐章。组合必须是高秩序性的过程。

上述这些观点在我们稍后去探索技术是怎样形成和发展时都会很有用。但是在此之前，我们依然需要回答第1章遗留下来的关于技术的一些更深层的问题：最初是什么使技术成为技术的？技术在最根本的本质上的存在是什么？或者换句话说，什么是技术的本质？是什么赋予一项具体技术以力量的？当然不太可能是它的原理，毕竟那只是一个理念。肯定是其他什么东西。

要解答这些问题，就要思考现象以及技术是如何利用现象的，而这正是我们下一章要做的事情。

THE

NATURE

OF

TECHNOLOGY

第3章

现 象

无论是简单还是复杂的技术，都是在应用一种或几种现象之后乔装打扮出来的。技术就是那些被捕获并使用的现象，是对现象有目的的编程。我们一直以为技术是科学的应用，但实际上却是技术引领着科学的发展。

考古学家们可以利用多种技术考证遗迹出现的时间。如果发现的遗迹是有机物，比如动物骸骨之类，就可以利用放射性碳定年法；如果发现了木质的残留物，比如一段柱子或过梁，可以应用树轮定年法；如果发现的是一个火坑，则可以选用古地磁定年法。

用放射性碳元素进行年代鉴定的原理是，活的有机体可以从空气中或通过食物链吸收碳，碳的内部会保有少量放射性同位素——碳-14，这种同位素会以恒定的速率一直衰减，直到无放射性的标准碳。生物体死亡后，就停止摄入碳了，这样在残骸中残存的碳-14开始规则性地衰减，出土时测量其碳-14含量，就能相当精确地确立它的年代了。

树轮定年法能够奏效是因为，树木年轮间的宽度会因降水量的不同而不同。所以在类似的气候或历史时期，树木的年轮会比较相似。通过与已知年代、地域相同的树木年轮样本比较，就可以精确地知道具有同种结构的树木生长的年代了。

古地磁定年法的作用机理则是，地球的磁场会以一种已知的方式随时间逐渐变换方向。火坑里的黏土或者其他材料在燃烧或冷却时都会留下一些弱磁性，和地磁联合起来考虑，就能提供关于这个火坑最后一次燃烧的时间。

还有另外一些技术可以达到类似目的。比如，钾-氩鉴定法、热发光定年法、水化层年代测定法、裂变径迹测定法等。但是我想提醒读者注意的是：这里的每一种方法都是依赖于一系列的自然现象而奏效的。

技术依赖现象是很普遍的。技术要达到某个目的，总是需要依赖于某种可被开发、利用的自然现象或自明之理。我说“总是”，理由很简单，一项技术如果什么都不开发，它就什么都不能获得。这是我提出的第三个原理。它和我之前所说的其他两个原理（组合和递归）同等重要。这个原理是说，如果你要审视技术，你总会在它的核心部分发现它所利用的效应。石油炼制要基于“气化原油过程中不同成分会在不同温度凝结”这个现象。一个下落的锤子则要依赖“动量传递”这一现象（即动量会从移动的锤子上传递到静止的钉子上）。

技术思想前沿

现象是技术赖以产生的必不可少的源泉。技术要达到某个目的，总是需要依赖于某种可被开发或利用的自然现象。无论是简单还是复杂的

技术，它们都应用了某一种或几种现象。

现象通常是显而易见的，但有时也很难察觉。我们面对一些非常熟悉的技术的时候，常常会碰到这种情况。比如，一辆卡车应用了什么现象？一辆卡车看起来似乎没有应用任何自然现象。然而，它确实用了一个，或者应该说，两个。一辆卡车的实质是一个平台，这个平台能够自动推进，同时比较容易移动。它自动推进的核心是基于这样一个现象：某种化学物质（例如，柴油）燃烧时会产生能量；而它能轻易移动的核心则基于圆的东西滚动起来比方的东西摩擦力要小（这一现象被理所当然地应用在轮子和轴承上）。后一种现象很难说是一种“自然规律”，它仅仅是一种可用的自然现象，而且看起来似乎很微不足道，但它一旦被开发出来，并应用于所有有轮子或能滚动的东西上的时候，它就变成了一种力量。

现象是技术赖以产生的必不可少的源泉。[1]所有的技术，无论多么简单或者多么复杂，实际上都是在应用了一种或几种现象之后乔装打扮出来的。

让我用下面的例子来说服你相信这个观点。假定你被要求测量某种不太容易测量的东西。比如你正身处零重力的外太空，而你需要测量一小块金属的质量。这时你不能将它放在天平上，或者把它悬挂成摆锤，或者挂在弹簧下端任其上下振荡，因为所有这些都需要重力。这时你可以把它挂在两个弹簧之间让它振动，或者想办法让它获得加速度，然后测量所需的力，这样你就可以间接测量它的质量了。但是请注意，这里你要寻找的是某个现象或者说某种效应，但是那个效应与你要测量质量这件事可能完全不是一回事。对于所有的技术，你需要一些可靠的现象来建立你的方法。

这样的例子听起来有点像传统本科生的物理试题。比如要求学生用一个气压计、一团线、一支铅笔以及封蜡去测量一栋建筑的高度。我可以举个类似物理试题的例子，它正是始于问题，再通过寻找相应的现象来解决问题的。这是一个现实的技术，实现它需要依赖四个基本现象。

题目是这样的：如何探察那些绕着遥远的恒星运行的行星（通常被称作系外行星）？这个技术是20世纪90年代由天文学家杰弗里·马西（Geoffrey Marcy）和保罗·巴特勒（Paul Butler）开发的。[2]因为那些行星太遥远了，直到最近，我们才有了可以直接观测它们的望远镜。而那时的天文学家则需要被迫去寻找证据以间接证明它们的存在。马西和

巴特勒从一个简单但细微的现象入手：假设恒星是飘浮在太空中的，它们之间有成千上万光年的距离，行星向它们所围绕的恒星提供微小的引力，这引起了恒星的规律的、重复性的振荡。恒星的晃动很轻微，每秒钟只移动几米。而且这种晃动过程实际上无比缓慢，因为它是在行星沿轨道移动数月，甚至数年的过程当中发生的。但是，如果天文学家能够探察到这些晃动，就能由此推断出一个行星的存在。

那么如何才能探察到这样细微的振荡（天文学家称之为“摄动”）呢？这里需要两个现象：一个是，从恒星发出的光能够被分解，形成光谱，因而呈现出连续的、颜色分明的色带（即光频）；另一个是，如果恒星朝我们的方向发生位移，这些线就会产生微小的偏移（即著名的多普勒效应）。马西和巴特勒将这两个现象结合起来，用望远镜锁定一颗恒星，分解它的光，形成光谱，并尝试去寻找谱线上数月或数年间发生的微小变化，从而检测恒星任何一次微小的摄动。这听起来简单，但做起来很困难。行星摄动在光谱上引起的变化太微小了，如果将光谱中一个特定的谱线比作音节中的中央C，那么马西和巴特勒正在寻找的将是从C到升C之间的一亿分之一的微小移动。如何才能观察到光谱中那么细微的变化呢？

马西和巴特勒用第四个现象来完成这个任务，这也是他们的最主要的贡献。他们让恒星的光穿过一个碘蒸气小室。当光穿过时，有特殊光谱特征的光会被气体吸收，这样一来，恒星光谱就呈现出像超市的条码一样的黑色的“吸收带”。这个碘蒸气过滤装置是不动的，所以这条带也不动，成为定尺。这样当恒星向观察者移近或远离时，光谱上的谱线的微小移动就可以被观察到。这项技术的实现需要不断地改进，马西和巴特勒用了9年的时间进行完善。尽管如此，为推断出外部行星存在而测量一个星星每秒钟不超过10米的位移——这项技术还是太不稳定了。

整个技术更像一个拼凑起来的物理实验，而不像一项高度精密的发明，事实也正是如此。这个例子阐明了如何应用现象的组合去达到目的。其中涉及了四个现象：行星的存在引起恒星摄动；恒星的光可以被分解形成光谱；如果恒星相对我们有位移，光谱线就会移动；光穿过气体可以产生一条固定的谱线作为记录恒星光谱中任何变动的基准。通过将上述现象加以合理地选取和组织，就形成了一项技术。之后，在这些效应的辅助下，天文学家们发现了150个新的系外行星。

我说过，技术是建立在概念或原理之上的，这和“技术是建立在现象之上”的说法一致吗？首先，原理和现象一样吗？回答是：不一样，至少

不同于我所指的“原理”。一项技术建立在某种原理、“某种方法”之上，这是一个技术过程得以开始的理念性基础。原理需要发掘出某个（或几个）现象来完成它。因此，原理和现象是不同的。比如某个特定的对象，如钟摆或石英晶体，会按照一个给定的频率摆动，这是一种现象。利用（using）这一现象来计时，便构成了一个原理，进而产生了时钟。高频无线信号遇到金属，会出现干扰和回声，这是一种现象。利用这种现象，通过发送信号然后接收回声来探察飞行物就构成一个原理，进而产生了雷达。现象只是简单的自然效应，因此独立于人类和技术而存在，现象本身并没有什么用。相比之下，原理是为达成某个目的而利用某个现象的理念（idea），它广泛存在于人类及“使用”的世界。

在实践中，现象在能够被应用于技术之前，一定要被“驯服”，并且作好恰当的准备。天然形式的现象很难被利用，需要巧妙的诱导，它们才能令人满意地运作起来，它们可能只在有限的条件下起作用，所以一定要建立正确的支持方式才能使它们为预设的目的服务。

> 技术思想前沿
>
> 现象在能够被应用于技术之前，一定要被“驯服”，并且作好恰当的准备。

这就是我在第2章说的支撑技术所起的作用。正如我所说，许多支撑技术是为基本原理服务的，它们为基本原理提供能量，同时管理和规范基本原理。但是也有许多支撑技术是为了支撑现象的应用以及安排它们去正确地实现需求而存在的。马西和巴特勒的碘蒸气室一定要准确控制在50摄氏度，这需要一组加热单元来帮助完成；恒星光谱在光谱仪中的分散非常细微，这需要计算机的加盟，以帮助校正；地球本身正在穿越空间，这需要进一步的以数据为基础的算法来加以调整；恒星的光会表现为密集的爆发，这需要更多的计算机算法去帮助筛选出真正的变化。所有这些需求都要求它们有自己的集成件和组件：加热单元需要绝缘和控温组件，计算机算法需要专门编写的软件等。因此，要使现象得以应用，需要大量的集成件和支撑技术。

这里的“模块”，与我在前面说的为了设计或者集装的方便而“组块”相比，有了更深的含义。子系统是技术，它们让实现特定目的所需的现象变得可用，同时子系统技术又有自己所依赖的现象。因此，一个实际的技术会包含许多现象一起作用。一个无线电接收器不仅是零部件的

集合，也不仅是一个信号加工厂的微缩版，而是现象聚集而成的交响乐——感应、电子间相互吸引和排斥、电子的发射、电阻电压下降、频率谐振……所有现象都被召集并组织在一起，为一场“音乐会”中的特定目的而工作。

总是有必要去设定那些现象，并且通过把现象分配给不同模块，从而把现象分开。我们不会在一个电子设备的内部将电容器放在感应器旁边（因为那样会导致不必要的振动）；我们也不会在飞机发动机里面的测量压力下降的装置旁边放置燃烧装置，我们会将这些现象分配到不同的模块中去。

技术的本质

相对于只将技术看作实现目的的手段，我们现在有了更直接的描述：技术是被捕捉到并被使用的现象，或者更准确地说，技术是那些被捕获并加以利用的现象的集合（a collection of phenomena）。我在这里用的是“捕捉”这个词，但是还有许多词可以用。我还可以说，现象是为了某种目的而被驾驭、控制、缚住、应用、采用、利用，或者开发的。然而在我的心中，“被捕捉并使用”是我认为最确切的表述。

> 技术思想前沿
>
> 从本质上看，技术是被捕获并加以利用的现象的集合，或者说，技术是对现象有目的的编程。

这种观察将我们带回到本书一开始提到的那个问题：技术的本质是什么？在最深的本质上，技术是什么呢？对我而言，答案就是我们刚才所说的：技术就是被捕获并使用的现象。或者反过来说，技术是被捕捉并被使用（put to use）的现象的集合。它之所以是核心所在，是因为一个技术的基本概念，即，使技术成为技术的东西，总是利用了某个或某些从现象中挖掘出来的核心效应。在本质上，技术是指向某种目的的，被编程了的现象。我这里特意用“被编程”这个词，是要强调使技术成为技术的现象是以一种有计划的方式被组织起来的，它们为“使用”这个目的而“共谱乐章”。

如此一来，我们可以用另外一种方式来表述技术的本质：技术是对现象的有目的的编程（programming）。

这个编程过程不一定十分明显。如果我们只想从技术外部审视技术，也没必要将这种内在的编程呈现出来。比如，当我们从外部看喷气式发动机时，我们能看到的只是它能够提供巨大的动力。在这个水平上，发动机只是一个手段，一个提供推力的装置。它可能是一件非凡的装置，但是我们仍然认为它不过是个装置。如果我们再深入一点儿，比如说为了维修，需要取下引擎罩，我们就可以看到它是一个由零部件构成的集合——一团管子、系统、电线、叶片等缠绕在一起。这时，这个电机已经是一个可见的可执行组合了。尽管这样看技术时，可能会令人印象深刻一点了，但是好像也就仅此而已。如果再继续深入一些，超出那些可见的部分，则会发现发动机真的是将一些现象“编程”在一起进行合作，是一场现象的“合奏”。

这里所涉及的现象没有一个是神秘难懂的，其中大多数甚至都是相当基础的物理现象。在此列举几种现象，括号内是那些偶尔使用它们的系统：当能量传递时，流体流速和压力可以改变（压缩机系统）；碳分子在高温和氧气混合时，会释放出能量（燃烧系统）；在源与汇之间大的温差会产生较高的热效率（燃烧系统）；不同器件之间的分子所形成的薄膜使它们能更容易地相互滑过（润滑系统）；流体撞击到移动着的表面可以产生“功”（涡轮系统）；荷载会使材料倾斜（特定的测压装置）；荷载可以由物理结构传递（负荷轴承和构件）；流体运动增加会引起电压下降（伯努利现象，用于流量测定仪器）；物体以一定速度被射出会产生大小相等、方向相反的作用（风扇和排气推进系统）。我们还可以加入更多的机械现象，进而我们也就可以增加更多的电器和电现象，使其成为电机控制、感应以及仪表系统工作的基础。继续列举下去，还可以增加光现象等。每个系统都会开发一些现象，每个系统都是整个装置的“分形”，因此现象的开发会不停进行下去，直至最基础的部分。

大量的现象被捕获之后，会被封存在各种各样的装置中，并被重复使用。这意味着在成千上万的零部件中，有些现象被成千上万次地重复使用着。所有这些现象被追踪、捕获、分类，然后恰到好处地在适当的温度、压力以及气流条件下产生作用；所有这些现象又恰好在适当的时刻，步调一致地产生作用。这些现象会忽略掉那些极端的振动、热量或者压力，而维持应有状态。当所有这些现象作用在一起，并产生了上千万磅的冲击力时，我们就不能再对其坐视不理了，因为这时，它们已变成了奇迹。

这样来看，操作过程中的技术（例如，喷气式发动机）不仅仅是工作中的一个物件，而且还是一个新陈代谢的过程。这不是一种令人熟悉的看待技术的方式。但是我的意思是，技术变成了一系列交互作用的过程，即一系列被捕获的现象的互相支持、互相利用、互相“交谈”，就如同计算机程序中的子程序之间的互相“呼应”（calling）一样。这里的呼应不一定需要像计算那样要有先后顺序，它可以是持续的、连续不断的交互作用。一些组件开启，另一些关闭，还有一些会持续工作；一些组件前后相继，另一些则并行不悖，还有一些只有在非正常情况下才会介入。

对于像飞机发动机这样的装置技术，这种呼应要求给定组件的运行要并行且持续，比如压缩系统需要在高压空气持续供给的条件下“召唤”或执行。这很像计算机算法中的并行操作，它们总在相互“交谈”，而且会持续互动。对于方法技术（例如，工业过程和算法），呼应则更倾向于前后相继。它更像标准的计算机顺序操作程序，但仍然是一种交互的动态过程，一种新陈代谢。

这样看，技术就不仅是手段而已。技术是为了某个目的而对现象进行的编程。技术是为了让我们能够使用而进行的现象的“合奏”。

> 技术思想前沿
>
> 基因是生物进化的基本单元，与此类似，我们可以把“现象”称为技术的“基因”。

这就有了一个推论。我在第一章曾说过，技术中没有纯粹的基因，但那并不意味着在技术进化的过程中完全没有和基因相似的东西。我认为，“现象”就是技术的“基因”。当然这种比喻并不十分贴切，但是这么想是很有帮助的。我们知道生物是通过激活基因来创造自身结构的，比如蛋白质、细胞、糖分等。以人类为例，人类大约有21 000种基因，而且在果蝇和人类之间，或者人类与大象之间，这个数字不会相差太多。单个基因与构成某种结构之间没有直接的关联，某一个基因是不能创造出眼睛，或者哪怕是眼睛的颜色的。现代生物学表明，创造了结构形状和形式的巨大变异的是基因组合，它们扮演了构成程序语言的元素。这个过程类似于将音调、节奏以及乐句结合起来，使之成为一种语言，从而创造出不同的音乐结构。生物性状和物种也是通过对那些几乎一样的基因加以“编程”，使它们以不同的顺序被激活而创造出来的。

技术也是如此。每个技术都是通过对一组固定的现象以不同的方式进行“编程”创造出来的。当然，随着时间的推移，新的现象，新的技术“基因”会不断加入进来。现象不是直接地被组合，它们首先被捕获并表达（expressed）为技术的元素，然后才能被组合。尽管比起生物基因来，可用的现象要少得多，但我们还是做了这样的类比：生物对基因加以编程从而产生无数的结构，技术对现象加以编程从而产生无数的应用。

有目的的系统

到此，有些读者可能会有些疑问，对此我要进行一些说明。我将技术定义为实现目的的手段，但是除了技术，还有许多可以被定义为实现目的的手段，但完全不像是技术的事情，比如商业组织、司法系统、货币系统或者合同，它们都可以被称为实现目的的手段。在这个定义下，它们也都是技术。进一步讲，它们的次生部分，例如组织或部门的分支，也是带有目的的手段，因此它们也具备了技术的性质。但是，总是感觉它们缺少某些“技术的本性”的东西，那怎么办呢？如果我们不接纳它们为技术，那么我的关于技术的定义就失效了，但是如果我们接纳它们为技术，我们则需要分析如何划定界限：马勒交响乐是技术吗？它也是实现目的的手段吗？比如说，它提供了一种感觉体验，而且它也有实现目的的组件。马勒是工程师吗？他的第二乐章（请原谅我这种双关用法）是为着某个目的而对现象进行的“合奏”吗？

我本来想通过把技术限定在设备和方法手段上来将这个问题回避掉，但是我现在不打算这样做。如果我们能习惯于把所有的手段，如货币体系、合同、交响乐、法律条文以及物理性的设施和方法，都看作是技术的话，那么与我已讨论的装置和原理相比，技术的逻辑就会适用于更加宽广的范畴，讨论范围也会因此扩大。

所以让我们先来看看为什么某些技术不太像标准的技术。货币是以交换为目的的手段，因此符合技术的定义要求（我这里谈论的是货币体系，而不是我们携带的硬币或者纸币）。它依据的原理是所有稀缺资源都可以作为交换的中介——黄金或者政府发放的纸币都可以，实在不行，香烟或者尼龙也行。货币系统利用了这样一种现象：只要我们相信其他人会相信一种交换媒介是有价值的，而且这种信任会在未来持续下去，我们就会相信这个系统。需要注意的是，这里的现象不是物理性

的，而是行为性的，这也就是为什么虽然“钱”满足了技术的定义，但却不太像技术，因为它不是基于一个物理现象。同样的理由可以解释我上面列出的那些不像技术的技术，也就是说，它们基于的现象是属于行为性的或制度性的，而非物理性的。

我们说雷达、发电机更像技术，是因为它们基于的是物理现象，而合同、立法系统等之所以不太像技术，则是因为它们基于的是那些非物理的现象——组织性现象或行为性现象，或者甚至是逻辑或数学的现象（例如，算法）。可以说，判别标准技术的标志就是看它是否是建立在物理现象之上的。

那么我们又该如何对待这些基于非物理现象的技术呢？为了本书的论述，我们当然可以承认它们都是技术，而且我们的确准备这样做。但同时我们也会承认，在日常生活中，它们确实通常不被看作是技术。马勒的一部交响乐只是一次审美体验而已，而一个软件公司也就只是一个组织。但是即使我们选择这样看待它们，我们也应该在心中牢记它们也是“技术”。马勒是在我们头脑中对现象进行了深思熟虑的“编程”：他在我们的耳蜗神经核、脑干、小脑、听觉皮层建立反射。因此，至少在这个意义上，马勒确实是一位工程师。

我们换一种说法，看看能否更简单明了地表达这个意思。如果仔细阅读，你会发现，本书一直讨论的是某类系统：我称之为目的性系统（purposed systems）。目的性系统是所有“实现目的的手段”的总体，既包括基于物理现象的技术，也包括基于非物理现象的技术。我们倾向于将某些手段，例如雷达、激光、核磁共振，称为传统技术，而将另一些手段，例如交响乐和组织，称为目的性系统。即便它们符合技术的特质，它们也更容易被理解为技术的“亲戚”。我们约定，大部分的讨论会针对狭义的物理技术，但是如果需要，我们也可以将范围扩展到非物理性的目的性系统。

> 技术思想前沿
>
> 目的性系统，是所有“实现目的的手段”的总体，既包括基于物理现象的技术，也包括基于非物理现象的技术。

上面的讨论貌似有点离题，但实际上它帮助我们确立了讨论的范围。我们认为，乐器、货币、法条、制度以及组织等，即使不依赖于物理现象，它们也确实都是手段或目的性系统，因此都应该在讨论的范围

之内。只要进行适当的调整，我之前提出的技术的逻辑对它们也同样适用。

捕捉现象

我在本章中提出，现象是所有技术的来源，技术的本质隐藏在为达成目的而去组织、协调现象的过程之中。接下来的问题是：在最初，我们是怎样揭示并驾驭现象的呢？

我认为，现象是隐秘的，如果不去发现或发掘，现象是不能显现的。当然它们也不是随机散乱分布的，而是团聚成各种相互关联的现象簇，这些现象簇各自形成不可见的效应——例如，光学现象、化学现象、电学现象、量子现象等——的“矿层”或“矿脉”，在漫长的时间中，现象簇里的效应一个一个被缓慢地、偶然地开采出来。越是浅层的效应，比如人类早期掌握的木头摩擦可以生热并生火的效应，是意外事件或偶然开发的结果；更深层的效应则像是矿层，比如化学反应，这其中某些效应的发现得益于早前的专家，但对这些效应的全面发现则需要系统的研究。对于隐藏最深的效应，比如核磁共振的量子效应、隧道效应或受激辐射等，则需要知识的积累和现代技术来揭示了。它们的发现需要现代性的发现或再现的方法，换句话说，它们需要现代科学的帮助。

那么，科学是如何揭示新效应的呢？当然科学不可能直接揭示新效应。从定义上来说，新效应就是未知的。科学往往是通过挑选不同于预期的事物来揭示效应的。比如某个研究人员在这注意到某些东西，某个实验室在那又忽视了某些东西，但之后又能发现那些细微线索。伦琴（Röntgen）是在操作克鲁克斯放电管（实际就是阴极射线管）时，发现几十厘米远的覆盖着氰亚铂酸钡的纸板微微地泛有红光，就这样偶然发现了X射线。有些时候，现象的蛛丝马迹是通过理论推理获得的。普朗克就是通过解释黑体辐射谱的理论探索，“发现”了量子（或者更准确地说，引入了能量可以被量子化的观念）。一个新效应的显现通常是作为某种尝试的副产品。例如作为基因技术的核心现象，DNA碱基配对互补的现象（即碱基A和T匹配，C和G匹配），就是沃森和克里克在尝试建构DNA的物理结构模型时的副产品。

技术思想前沿

揭示新现象有三种途径：

- 重新关注在实验过程中被忽略的细节。
- 通过理论与推理寻找现象的蛛丝马迹。
- 某种尝试的副产品。

如此看来，新效应的发现好像是逐步地、各自独立地进行的。但实际上并不尽然。通常，一个现象簇内已知的效应会导致后面效应的发现。1831年，法拉第发现了电磁感应现象。[3]他先用铜丝缠住一个铁环的半边，然后连上电池，这样在铁环内就产生了一个强大的磁场。然后他将铁环的另一边也用铜丝缠上，当他转换磁场时，会在第二个线圈中引起或“感应”出电流。他用磁罗盘检测后发现：“磁针会立即敏感地产生反应。”同样，断开电源时，磁针也会偏转。法拉第据此发现一个变化的磁场会在附近的导体中产生电流。

我们可以说法拉第应用了科学的洞察力和实验完成了这个发现，事实也确实如此。但是，如果我们再深入一点，就会看到，法拉第发现、理解感应现象应用的仪器是建立在以前的发现基础之上的。换句话说，他所使用的电池应用的是以前发现的电化学现象；线圈缠绕在铁之类的磁性材料上会产生强磁场的现象是由荷兰乌特勒支城的格里特·莫尔（Gerrit　Moll）此前刚刚发现的；而他用的电流探测仪的原理则源于汉斯·奥斯特11年前关于电流会使磁针偏转的发现。总之，法拉第的发现之所以成为可能，是因为此前已有相关的现象被理解和驯服了。

对于现象，这种情况很普遍。随着一个现象簇被不断发掘，先前发现的效应为此后效应的发现提供了方法和启示。一个效应引发另一个，然后是下一个，直到最后相关现象的整个矿脉都被发掘出来。一个现象簇形成了一个井下硐室，硐室之间由矿层或巷道相连，一个通往另一个。然而这还不是全部，一个硐室，一个现象簇，会通过巷道通往任何硐室，即使通向完全不同的现象簇也可以。比如，先前的电现象就导致了后来量子现象的发现。现象构成一个被发掘出来的相互联系的硐室和巷道系统，而且整个地下系统都是可以相互连通的。

一旦现象被挖掘了，它们又是如何被转化成技术的呢？这是我们后面要详细探讨的问题。现在，让我仅就其本身的性质进行一下说明。现象在本性上就是可以做点什么。当人们意识到这个潜在用途时，就可以驯服它做事了，比如将变换磁场感应电流这一现象转化为发电的手段，两者之间其实并不特别遥远。

当然不是所有的现象都能够被驯服利用，但是，一旦一个现象簇被发现，就会有一连串的技术随之而来。1750　　—1875年，主要的电现象，例如，静电效应、电蚀作用、由电场和磁场导致的电流偏转、感应现象、电磁辐射，以及辉光放电现象都被发现了。随着对这些现象的捕捉和驯服，随之而来的是一系列的方法、工艺及设备，其中包括电池、电容和电感、变压器、电报、发电机和电动机、电话、无线电报、阴极射线管、真空管等。

技术思想前沿

现象是以累积式建构起来的。现象首先被捕获，然后这现象被用于制造设备，并接着进一步揭示新现象。

当先前的现象在仪器设备等的共同帮助下揭示了后面的现象时，这种累积式的建构过程就缓慢地发生了。如此这般，现象不断地将自身揭示出来，步步为营、累积式地前进——首先现象被俘获，然后应用这些现象制造设备，接着进一步揭示新现象。

技术与科学

当然，这一切的发生都需要理解，或者至少需要知道关于效应以及如何利用这些效应的知识。那么，这种知识在哪里呢？很明显，知识，即信息、事实、真理、一般原理，对我们正在关注的这个主题至关重要。让我以尖锐的方式再问一遍同样的问题：现象的正式知识（科学）在哪里呢？科学又是如何与现象的使用联系起来的？或者可以直接问，科学是如何联系到技术本身的？

技术思想前沿

科学与现象的关系：

- 科学提供观察现象的手段。
- 科学提供与现象打交道时所需的知识。
- 科学提供预测现象如何作用的理论。
- 科学提供捕获现象、为我所用的方法。

科学对于发现现代现象以及那些隐藏更深的效应集群，并据此建构起技术来说，无疑是必要的。它提供观察现象的手段，与它们打交道时所需的知识，预测它们将如何作用的理论，以及捕获它们为我所用的方法。因此，在我们和现代现象打交道的过程中，科学的介入很有必要。这样讲没什么不对，甚至是无法反驳的，但是接下来的推论则一直是学术界具有争议性的焦点：科学发现新现象，而技术则利用了它们，所以看起来科学在发现，而技术则在进行应用。

技术就是科学的应用吗？这种观点当然有它的支持者。已故的杰出工程教授约翰·曲克索（John G. Truxal）就曾宣称："技术，就是应用科学知识来达到人类特定的目标。"[4]

那么，技术真的就是科学的应用吗？仅此而已？我认为不是。至少真相要比这复杂得多。过去许多技术（例如，动力飞行）的诞生都几乎与科学毫无关系。事实上，直到19世纪中期，技术才开始大规模地向科学进行"借贷"。科学之所以在这个时候链接到技术，不仅是因为它能对结果提供更多的洞见和更好的预测，还因为一些新的现象簇开始被揭示了，比如，电学和化学的现象簇，而就它们的规模或所处的世界而言，如果不借助科学的方法和仪器，人类就无法直接观察到。在拿破仑时代，要建筑一个传递信息的巨大的木制双臂信号装置，常识就已足够。但是如果想创造一种用电进行信息传输的方法（类似电报），则需要关于电现象的系统知识。技术之所以应用科学，是因为这是去理解深层现象的工作机理的唯一的方式。

但是这并不意味着技术人员只是向科学思想伸手，然后简单地对其加以应用。技术人员应用科学思想就如同政客们应用已故的政治哲学家们的思想一样，他们日复一日地使用这些思想，对其起源的细节却知之甚少。但这并不是出于无知，而是因为起源于科学的那些思想会随着时间的推移被消化吸收到技术体自身当中，例如融入电子工业或生物技术。它们在这些领域中与经验和特定的应用融合在一起，创造出它们自己固有的理论和实践。所以，断言技术只是科学的"应用"是幼稚的，毋宁说技术是从科学和自己的经验两个方面建立起来的。这两个方面堆积在一起，并且随着这一切的发生，科学会有机地成为技术的一部分，被深深地融入了技术。

技术也同样被深深地融入了科学。科学需要通过观察和推理，来获得其洞见，但正是对方法和设备的利用才使得观察和推理变得可能。科学可不是在很小的程度上利用仪器和方法——利用技术——来观察自然

界。在开创现代天文科学方面，望远镜与哥白尼和牛顿的推理同等重要。如果没有X射线衍射的方法、设备以及提取和纯化DNA所必需的生化方法，沃森和克里克也不可能发现DNA的结构（以及其后的互补碱基对现象）。没有观察和理解现象的仪器，没有显微镜、化学分析方法、光谱学、云室、测量磁电现象的仪器、X射线衍射分析方法以及大量衍生方法，现代科学根本不可能存在。虽然我们通常不这样看，但上述所有这些其实都是技术，科学正是从这里出发并建立起它的理论的。

那么科学实验呢？它们也和技术有关联吗？当然有些实验只是寄希望于幸运的发现而进行探索。但是严肃的科学实验一定系统地探索大自然的运作，并且在头脑中总带有一个明确的目标。因此，它们是实现人类目标的工具，它们是技术意义上的方法，通常包含或体现在物理仪器之中。当罗伯特·密立根（Robert　Millikan）在1910年和1911年实施他那著名的油滴实验时，他的目的是测量一个电子的电量。他建构（组合）了一个方法来完成这件事。这个理念，或曰基本概念，就是将很少数量的电子附着在很微小的油滴上，使用已知强度的电场，控制带电油滴的运动（带电粒子在电场里会被吸引或排斥）。在重力和电场力交替作用下，让油滴反复上升、下落若干次。通过测量油滴在已知电场（并单独算出它的大小）的运动，密立根就能计算出电子所带的电量。

密立根用了5年多的时间来不断完善实验方案，以不同的方式尝试改进实验方法。[5]在最终选定的测量方式当中，他选择用喷雾器喷出微小的油滴，然后将负电荷（电子）附着在这些油滴上面。这样他可以通过显微镜挑选出特定的油滴，让它在重力（没有磁场）作用下在两条水平十字准线之间自由下落，测定油滴在空气中坠落的时间，再运用斯托克斯流体运动方程最终确定电荷的大小。然后再切换成正电场，使带负电荷的雾滴向上飘移或静止不动，达到和重力制衡的效果。由油滴的运动速度（或必要的控制电压）及大小，就可以计算出它的电荷。密立根以这种方式观察了几十个油滴，发现油滴所带的电量总是某一个最小固定值的整数倍，并和附着在油滴上的电子数相对应。密立根因而得出结论，最小的整数值就是一个电子的电量。

像所有优秀的实验一样，密立根的实验优雅、简单。但是我们要注意它的实质所在，即它是达到目的的一种手段，是一种技法（technique）或方法技术（如果你注意到其中的步骤的话）。这些步骤发生在各个组成部分（喷雾器、油滴室、带电板、电池、离子源、显微镜）的建构过程中。而这些组成部分本身也是实现目标的手段。事实

上，在密立根的工作中，最引人瞩目的是他倾注在建构实验方法上的努力。因此，他为之奋斗了5年时间的任务，他对测量方式的完善，实际上是技术性的：他需要不断地、一件一件地完善各个零部件，即那些构成实验装置的集成配件。密立根本来是在探究一种现象，这完全是科学的方式，但是为完成这种探究，他建构的却是一个方法——一种技术。

这正是科学探索的方式。它使用仪器和实验等技术形式来回答特定的问题。但是，科学的自我建构当然不止这些。它用科学的解释、推理和理论建立了自己对世界如何运作的理解。我们确实可以说，至少从这些方面来看，科学是远离技术的。

但也不尽然。解释、说明这类事当然不太像是技术，但它们是带有一个目标的建构。它们的目标是弄清楚这个世界上被观察到的一些特征是如何运作的，它们的内容是一些依规则组合在一起的概念性要素。所有的解释都是从相对简单的部分建构起来的。当牛顿“解释”行星轨道时，他建构了一个概念的版本，这个概念版本从质量、质量间的重力作用以及它的运动定律等更简单的部分开始建立起来。也可以说，他建构了一个关于行星如何绕轨运行的数学“故事”。这个故事是从已有内容和规则的基础上讲起的。由此我们甚至可以将牛顿的理论，或其他解释性理论也称作技术了。如此一来，我们便乐见，科学也是目的性系统，它拥有技术的形式，或者至少可以说它是技术的堂兄弟。

由此可见，科学不仅利用技术，而且是从技术当中建构自身的。科学的这种自身建构当然不是来自于桥梁、钢铁或运输这些标准技术，而是从仪器、方法、实验以及它所采用的概念中而来的。这没什么可惊奇的，毕竟科学是一种方法：一种关于理解、探究、解释的方法，一种包含许多次级方法的方法。回到它的核心结构（core structure），科学就是一种技术形式。[6]

最后的这个观点可能会干扰甚至吓到我的读者。所以，我需要澄清一下我在说什么和我没在说什么。我现在说的是，科学形成（forms from）于技术，即仪器、方法、实验和解释等，它们是科学的肉身。我不是说，科学和技术是一样的。科学是内里孕育着美的事物，一种超凡脱俗的美，而且它的内涵要比由各种仪器、实验、解释构成的那个核心结构多得多。科学是一系列教寓观念，即自然在本质上是可知的，可以被探察、被究因的。如果以高度控制的方式对现象及其背后的含义进行探索，就可以获得对自然的理解；科学是一套实践和思维方式，包括理论化，想象和猜测；科学是一系列认识（knowings），一系列由过去的

观察与思考积累起来的理解；科学是一种文化，一种关于信仰与实践、友谊与思想交流、观点与信念、竞争与互助的文化。

> 技术思想前沿
>
> 科学和技术是两个不同的概念。科学建构于技术，而技术是从科学和自身经验两个方面建立起来的。科学和技术以一种共生方式进化着，每一方都参与了另一方的创造，一方接受、吸收、使用着另一方。两者混杂在一起，不可分离，彼此依赖。

这些都无法简化为标准技术。事实上，可以设想一种没有技术的科学，没有望远镜、显微镜、计算机、测量仪器，这样的科学只建立在思想和猜想之上。但这种设想更证实了我的观点，即没有技术的科学将是软弱的，这样的科学和希腊时期的思辨科学相比并没有什么不同。

那么这又提示了我们什么呢？我们可以这样说，技术是对现象的驾驭，而这很大程度上是由科学揭示的。同样，科学也建构于（builds from）技术，或者说，科学是从它的技术中形成的，从那些要使用技术的仪器、方法和实验当中形成的。科学和技术以一种共生关系进化着，每一方都参与了另一方的创造，一方接受、吸收、使用着另一方。这样做之后，两者混杂在一起，不可分离、彼此依赖。科学对于揭示和理解深藏的现象至关重要，而技术对于促进科学也同样重要。

最近，经济史学家乔尔·莫基尔（Joel Mokyr）就提出，技术是随着人类知识的增长而推进的。他是这样阐述的："过去400年呕心沥血的知识积累，辅之以社会的、科学机制的推动及知识扩散，共同奠定了工业革命和现代技术的基础。"[7]

我相信这一点。事实上，我和莫基尔都相信，无论怎样讲都不会夸大知识的重要性。你不可能在缺乏气流流过翼面（这里指涡轮机和压气机的叶片）的知识的情况下设计出现代的喷气式发动机。不过，我想表达一下与莫基尔的观察不一样的意见。因为在现象被发现和探索初期，围绕在它周围的通常是对这些现象的理解的半影区(7)。这些对技术发展有很大帮助的对现象的理解（理论和知识）都来自于这个半影区。事实上，直到现在，理解的半影区也是不可或缺的。曾经常识可能直接产生新的设备，比如纺织机；而现在只有详细、系统、可编码的理论知识才可能产生基因工程或微波传输等新技术，或是帮助人们寻找太阳系以外的行星。这是由于现代技术所使用的现象是无法仅仅通过随机观察和诉

诸常识就能被发现和理解。不过，这只是故事的一半。新的设备和方法形成于被发现和理解的现象，技术反过来会帮助建构进一步的知识和理解，以及帮助揭示更进一步的现象。知识和技术就是以这样的方式一起聚集发展的。

我已经在本章讨论了现象以及技术如何从现象中产生，但这对于技术会随时间发展意味着什么呢？

> 技术思想前沿
>
> 新现象与新技术构成一个良性循环。新现象提供了发现新现象的新技术，或者说新技术发现了导致新技术的新现象。

大自然中存在着许多现象簇，几万年来，我们已经有目的地挖掘了这些现象：史前对火的利用和金属加工、17世纪的化学和光学、18和19世纪的电、20世纪的量子现象，以及20世纪末的遗传效应。其中许多效应已被驯服，并被运用到技术当中了。同时，它们又成为建构未来技术的潜在模块。而有些技术（例如，科学仪器和方法）则主要用于帮助发现新现象。这是一个良性的因果循环。我们可以说，新现象提供了发现新现象的新技术，或者说新技术发现了导致新技术的新现象。无论哪种方式，技术和已知现象的聚集是前后相继地发展的。

这一切都不应被理解为技术始终是直接产生于现象的。大多数创造技术的现象模块与驾驭一个现象之间的距离可不只是几步之遥。火星探测器是由驱动马达、数字式电路、通信系统、转向伺服系统、摄像头和轮子组合起来的，但是对以上每样东西背后的现象的直接认识并不能直接促成火星探测器这个技术的完成。大多数技术都是如此。但是要谨记，所有的技术，包括行星车，归根到底都来自现象的。所有的技术最终都是现象的“合奏”。

现在这一点应该已经很明确了，没有现象，技术就不会存在，但反之则不然。现象本身与技术无关，它们（至少物理现象）仅仅是存在于世界中，我们无法控制其形成或存在。我们所能做的，只是在某些地方利用它们。如果我们这个物种在另一个拥有不同现象的宇宙诞生的话，我们将开发出不同的技术；如果我们过去发现现象的历史序列有所不同的话，我们也将开发出不同的技术。假设在宇宙的某个地方，我们认为正常的现象不再起作用了，那么，现有技术将失灵。我们对宇宙的了解只能通过追寻现象提供的暗示来逐步增加。这听起来像是一出科幻剧，

但是用不着远离地球，这种失灵就可能发生。在太空中，连最简单的事情（例如，喝水）都必须重新考虑，而我们可能仅仅缺失了一个效应：引力。

THE NATURE OF TECHNOLOGY

第4章 域

为了共享现象族群和共同目标，或者为了分享同一个理论，个体技术就会聚集成群。这种集群就形成了“域”。工程设计是从选择某个域开始的，这个自动和下意识的选择过程叫作“域定”。设计工作就像是用某种语言所进行的写作或表达。

当现象簇（包括化学的、电子的、量子的）被开发并加以利用以后，就自然而然地引起了技术的聚集。比如，与电以及电的效应有关的设备和方法（例如，电容器、感应器、晶体管、运算放大器）会自然而然地聚集到电学中来。它们和电子介质结合之后，彼此就可以方便地“对话”。光以及与光传播相关的元素（例如，激光、光纤、光学放大器、光学传感器等）聚集成光电子学。它们互相传递光量子和光能单位，以服务于不同的操作。这样的聚集构成了我在第2章所谈到的复数技术。现在我还想就这个问题再展开谈一下。我认为每一个“集群”都形成了一种语言，在这种语言中，一些具体的技术（设备和方法）在这种语言里作为表达聚集在一起。

为什么个体技术要聚集成群？原因之一是它们是共享了效应或现象簇。但这不是唯一的理由，技术聚集在一起还因为它们分享了共同的目标：斜拉桥的钢索需要锚固设施，锚固设施需要重型螺栓，这样一来，钢索、锚、重型螺栓就自然集成在一起了。那些分享共同物理强度和规模特征的要素，大梁、桁梁、柱子、钢梁、水泥板等聚集在一起，它们在强度、尺度和应用范围上是匹配的，它们因此成了结构工程的构件（building　blocks）。就这样，要素积聚成群，进而为形成可用的次级构件服务。基因工程方法，即DNA纯化、DNA扩大、放射性标记术、排序、通过限制性内切酶进行切割、克隆基因片段以及表达基因筛检，组成了一个天然的装满模块的工具箱，特定的程序就从这个工具箱中产生出来。

有时候，要素聚集的原因是它们可以分享同一个理论。例如，用来概括和分析数据并实施统计性测试的统计软件包，拥有一个共同的假设前提，即抽样在总体上应是正态分布的，因而统计软件包工作的基础是共享了所操作数据在这个性质上的假设。描绘一个技术集群的通常是某种形成的共性，即某些可以使共同工作成为可能的、共有的、自然的能力。

我称这种集群、这种技术体为域。一个域可能是用来产生设备和方法的要素的任意集聚，以及产生这些设备和方法所必需的实践、知识、组合规则及思维方式等的集合。[1]

技术思想前沿

> 某种形式的共性，或者是可以使共同工作成为可能而共同固有的能力，可以定义为一个技术集群，对于这种集群或技术体，我们称之为域。

对我们来说，如果想弄清楚技术是如何形成和进化的，关键是要保持个体技术和域之间的清晰界限。但是有时候，这个界限看起来是模糊的，比如雷达（个体系统）和雷达技术（工程实践）听上去非常相像，但其实不然。

一项技术（个体技术）是要完成一项工作，达成一个目标的，而且这个目标经常非常特殊。而一个域（复数的技术）则不需要完成工作；它仅仅是以一个工具箱的方式存在，你可以从中选取有用的元器件或一系列的应用规范。一项技术界定一个产品或一种工艺。一个域不界定任何产品，但它构成了一群技术——一组互相支撑的装置。当这组互相支撑的装置由生产它们的公司来表现时，它就界定了一个产业。一项技术是被发明的，它是由某个人集结组装的；一个域（例如，无线电工程）则不是被发明出来的，而是从它的个体的组件开始一点一点展现而成的。一项技术（例如，一台电脑）只向占有它的人展示某种潜力；一个域（例如，数字技术）则将潜力赋予了整个经济，并及时地将其转化为财富和政治权力。

技术和域还有另外一个不同之处。技术所占据的是模块、次级模块和零部件这样的层级关系，而域所占据的则是次级域、次次级域。电子学包含电子模拟技术和数字电子技术两个次级域，而次级域又包含着次次级域，比如固态半导体元件，在次次级域中又包含了砷化镓和硅器件这样的次次次级域。

对上述差异的总体表述是：单个技术之于技术体，就如同程序之于编程语言一样。

域定

工程中的设计是从选择一个域开始的，也就是说，是通过选择合适的元件群来构建一个设备。当建筑师设计一个新的办公大楼时，从视觉和结构上考虑，他们可能选择玻璃–钢架的组合而不是花岗岩，他们会在含有不同特性材质的“调色板”中进行选择，我们将这种“调色板”也称为域，而将这个选择过程称为“域定”（domaining）。域定过程常常是

自动的、下意识的。一位航海雷达的设计者会在电子元器件中毫不犹豫地“域定”主控振荡器，仅仅是因为没有更适合的域了。

但有时选择也很费思量，比如，一位设计者如果要集成计算机操作系统，那么他就需要在Linux系统和Windows系统之间进行选择。当然，任何规模较大的技术通常都需要几个域的同时加入，比如，一座发电站是从建筑施工、水利、重型电机和电子等不同的域中选取集成而来的。

在艺术领域，域的选定在很大程度上关乎风格。作曲家会让一个主题穿梭于管弦乐中不同的域——或管乐部分或弦乐部分，以此来获得感受和对比。在技术领域上，域的选定不取决于情绪或感受，而取决于它所能完成的集成的便利程度和效率，以及它和其他集成形成链接的容易程度以及成本。技术中的域定通常是很实际的。

> 技术思想前沿
>
> 工程设计是从选择一个域开始的，也就是要选择一组适合建构一个装置的元器件，这个选择过程，我们称之为“域定”。

针对给定目的的“域”的选择是随着时间的推移而变化的。在数字化技术出现之前，飞机设计师们是在机械和液压技术的范围内选择控制机翼和稳定器表面系统的。他们用推杆、拉杆、电缆、滑轮和其他机械链接将飞行员的操纵杆和那个用来移动飞行操纵面的液压机械舵联结起来。20世纪70年代，他们开始以数字化的方式在一种新式的、被称作电传操纵（fly-by-wire）的新技术体当中重新域定飞机的控制系统。新系统捕捉到飞行员的行动和飞机目前的运动状态，然后以信号的形式通过电线发送到计算机中进行处理，最后计算机再次通过电线将必要的调整信号传输到用来控制飞机操纵面的快速液压传动装置上。

电传操纵的飞机控制系统[2]要比之前使用的机械轻得多，且可靠性更高，反应也更快，新的控制系统能够在毫秒之间对变化作出反应，而且它们是“智能的”。通过计算机控制的电传操纵比人更精确，甚至可以纠正飞行员的不佳决定。

事实上，新的域使得新一代“内在不稳定”（inherently unstable，用技术术语讲就是“放宽静稳定”）的军用飞机设计成为可能。这种设计赋予军用飞机一种优势。就像控制一架不稳定的双轮自行车要比控制一

台三轮车更容易一样，控制一架内部不稳定的飞机要比控制内部稳定的飞机更容易。新的控制系统使飞机平稳的方法和骑车的人通过反向运动来抵消不稳定性以保持自行车平稳的方法是一样的。人类飞行员无法如此迅速地作出反应，因而在早期的手动技术中，内在不稳定飞机是无法飞行的。

> 技术思想前沿
>
> 重新域定（redomained），是指以一套不同的内容来表达既定的目的。重新域定不仅提供了一套新的、更有效的实现目的的方法，还提供了新的可能性。这意味着技术的颠覆性改变。

因此，我们可以说飞机控制系统经历了一场“创新”，并且他们真的做到了。但是我们应该更准确地说，飞机控制系统是以不同的方式被域定了，或者说是被重新域定了。这种区分很重要。历史上的创新常常是在已有的技术上进行的改进，例如，一种更好的建造圆屋顶的方法，或者一种更有效率的蒸汽机。但是显著的改进实际上是“重新域定”，即以一套不同的内容来表达既定的目的，就像动力供给的方式从水车技术发展为蒸汽技术那样。

重新域定之所以强大，不仅仅在于它们提供了一套新的、更有效的实现目的的方法，更在于它们提供了新的可能性。20世纪30年代，人们可以用由水泥制造的、4米多或更高的巨大的声反射镜去监察跨越海峡飞往英格兰的飞机。它需要搜集来自32公里远的声音，然后交由具有超敏感听力的人进行处理。到第二次世界大战爆发时，雷达被用来实现同样的目的。雷达的有效范围更大。我们可以说，飞行侦查任务采用（或者更确切地说是引发创造）了雷达技术，但是我更愿意说是这项技术利用了一种新的、更强大的域，即无线电工程。这个域横扫了整个世界，其组件后来又被发掘出许多用途。

域内发生的某些变化是技术进步的主要方式。

当一个新的域出现的时候，它可能并不能直接显现明显的重要性。比如，最初无线电仅限于电报通讯，即从船和岸之间用无线电方式发送信息。但是，新域的出现不仅在于它能做什么，更在于它能带来的新的潜能。这种潜能会激发当时技术专家的灵感。1821年，查尔斯·巴贝奇（Charles Babbage）[3]和他的朋友，天文学家约翰·赫歇耳（John Herschel）正在为天文学会准备一张数学表。两人之前比较了同样一张

表，然后分头计算。巴贝奇因为错误频出而绝望地说："我真希望上帝能够用蒸汽来进行这些计算。"巴贝奇的话在今天听起来很奇怪。后来，他设计了一个计算设备，但不是用蒸汽，而是用手柄和杠杆驱动的。这里应该注意到的是，他诉求的不是新的设备，而是新的域！让我们向那个时代的奇迹之域致敬吧，在19世纪20年代，蒸汽定义的是一个关于可能性的新世界。

实际上，域不仅定义可能性，而且可以定义一个时期的风格。回想一下，在维多利亚晚期，科幻小说中对宇宙飞船航行的憧憬，就可以明白我在说什么。我这里想到的是儒勒·凡尔纳（Jules Verne）的《从地球到月球》中的场景[4]，其中折射了19世纪60年代的法国映像。如果你看到这本书中描绘的航天器以及它的着陆装置，就能体会到，那些辨别出它们所属时代的因素，不是其外形或设计（这是航天器独有的），而是因为它们所选用的组件是18世纪中叶才可能出现的技术。这些技术包括铸铁的太空船、用火炮抛送太空船，整个冒险发生在砖和熟铁的建筑结构当中。这样的技术组件和使用方式不仅反映，而且定义了一个时代的风格。

时代创造着技术，技术同时也创造着时代。技术的历史不仅是单个发明和技术的编年史，比如印刷机、蒸汽机、贝塞麦炼钢工艺、无线电、计算机，它也是时代（整段时期）的编年史，通过确定整个时代各种目标的聚合来定义时代。

当你步入一个被精心设计的博物馆，你就可以看到，或者应该说可以感受到这一点。博物馆通常会有一个或者几个特殊的房间用来回顾这些历史上的特殊时刻。洛斯阿拉莫斯历史博物馆就用这样的展览，来展现战争年代的人工制品和技术。你可以看到化学用的曲颈瓶、计算尺、曼哈顿计划的身份证、身着宪兵警察制服的假人、汽油配给的卡片、旧卡车和吉普车的场景等，所有这些都与当时完成任务的典型的手段，也就是技术有关。而这些技术比任何其他东西都更能定义它们所描述的时代。洛斯阿拉莫斯展览本身是小规模的，只有两个或三个房间，但进入它，就相当于进入了那个时代。

域不仅定义时代，它们也定义时代的边界。想象一下，如果用巴贝奇时代的域来描绘一幅地震数据地质图会怎样。它可能需要巴贝奇亲自设计设备，组装一台地震波分析仪，并依据爆炸回声进行地图的渲染。这个机器将是一个奇迹，它也许会带有一个犄角形的听筒，一个辅助的铜齿轮和杠杆，以及转盘和着墨的绘图笔。这个设备工作起来缓慢、复

杂，而且它将专门用来进行地震勘测。巴贝奇时代的这些域（机器、铁路、早期工业化学）的触角很短，因此所能提供的可能性很有限。今天，我们拥有的域则可以提供更广泛的可能性。实际上，一个域的有效性一半来自于它的范围，即它能开启的可能性，另一半则来自于它能否为不同目的进行反复的、相似的组合。这有点像过去的排字工人都会在手边事先准备好一些常用字的组合（法国18世纪的印刷工人称它们为“陈词滥调”），以便使用起来更加方便、高效。

设计就如同语言表达

一个新的设备或方法是由一个域中适用的零部件，或者也可以说是适当的词汇聚集而成的。从这个意义上来看，一个域构成了一种语言，当某个域在产生一件新的技术产品时，就是这个域在以某种语言进行表达。这使得技术在整体上如同多种语言的集合，因为每一项新技术都可能从多个域中汲取元素。这也意味着技术中的主要活动，即工程设计，变成了一种写作方式，一种（或几种）语言的表达。

> 技术思想前沿
>
> 一个域就相当于一种语言，当某个域在产生一件新的艺术品时，就相当于这个域在以某种语言进行表达。语言的组织必须依据语言规则，设计的建构也要根据域允许的组合规则进行，这种规则就是语法。

我们并不熟悉这种看待设计的方式，但是它值得我们考虑一下。语言中有清晰和不清晰的表达方式、恰当和不恰当的语言选择之分，设计亦是如此；语言可以简明扼要，设计亦是如此；语言有不同程度的复合句式，设计亦是如此；用语言表达一个理念可以只用一个简单句，也可以用一整本书，并辅之以若干支撑材料来呈现一个主题，设计亦是如此。语言中任何目的的表达都可以有很多选择。类似地，技术为达到任何目的也可以选择多种组合。如同语言的组织必须依据语言规则一样，设计的建构也要根据域允许的组合规则来进行。

我将这种规则称为语法。可以将语法想象成一个域的“化学”规则，即确定什么可以被允许进行组合的一组原则。这里用的“语法”就是我们常用的那个意思。亨利·詹姆斯（Henry James）曾谈及“绘画语法”[5]；生物化学家埃尔文·查戈夫（Erwin Chargaff）1949年在论及他在DNA化

学方面的发现时就曾说过："在眼前模糊黑暗的轮廓中，我开始看到了生物学的语法。"[6]詹姆斯和查戈夫所指的语法，并不是绘画或分子生物学的特性，而是指绘画元素或生物元素之间相互联系、相互影响、相互结合并形成结构的方式。

一个域的语法决定它的元素如何被组装在一起，以及在什么情况下它们会结合在一起。它决定什么东西在"起作用"。从这个意义上讲，我们就有了电子学、水力学以及基因工程学的语法，对应更精细的域，还会有次级语法和次次级语法。

这样的语法是从何而来的呢？[7]当然，毫无疑问，它最终一定是来源于自然。电子学的语法背后是电子运动的物理学以及电现象的规律。DNA操作的语法背后是核苷酸和与DNA一起工作的酶的内在特性。语法在很大程度上反映了我们对"特定域中自然是如何工作的"这个问题的理解。这种理解不仅是来自于理论[8]，也来自于经验积累：工作温度和压力参数、机器配置和工具的选用、过程计时、材料强度、集成件间的安全距离——无数的知识碎片组成了各种技艺的"烹饪术"。

有时这种知识可以还原为拇指法则(8)（rules of thumb）。飞机设计界早就从多年的经验中知道："成功的喷气飞机的引擎推力重量与加载的飞机之间的比重永远都大约介于0.2～0.3。"但更多时候，知识通常不能转化成这样的法则。在这方面，技术和艺术实际上没什么不同。正如巴黎蓝绶带烹饪（Cordon Bleu）(9)不能简化为一套书面原理，专业电子设计也不能。无论是烹饪还是工程的语法，都不仅作为规则存在，还必须作为一套被认为是理所当然的潜经验（unspoken practices）存在，因为实践知识（practical knowing）可能没办法用语言进行充分的表达。语法由文化经验和应用技巧构成。如此一来，语法就不仅存在于使用者的头脑中，不仅出现在课本中，还存在于它们共享的文化中。它们出现时可能被当作规则，但是最终则会成为一种概念化的技术、一种思维的方式。

事实上，正如口语表达的清晰度不仅取决于语法一样（它依赖于词汇的深层含义及其文化相关性），技术的清晰表达也不会仅仅依赖于语法。技术的清晰表达需要相关域的深层知识，包括：所使用组件的词汇的流利程度；对标准模块、以前的设计、标准材料、相关技术的熟悉度；一种关于什么是自然的，什么会被该领域的文化所接受的"洞悉"（knowingness）；直观知识、横向沟通、感觉、曾经使用的经验、想象力、品位——所有这些都是有价值的。

技术专家詹姆斯·纽科姆（James　Newcomb）在谈到提供节能服务时说："做好这种业务需要数以千计的个体技术知识，以及为特定目的吸收和优化组合这些技术的能力，同时要考虑到交互影响、控制系统、过程影响、能源经济和节约需求的差别。这要求厨艺大师级那样的技艺，而不是杂货店采购员的水准。"[9]

好的设计事实上就像一首好诗。这不是指诗的崇高感，而是指从众多的可能性中为每个部分作出完全正确的选择。每个部分必须紧密配合，各部分的运行一定要准确，必须符合与其余部分的互动规则。一个好的设计[10]的美感在于合宜性，在于为所获得的东西付出最少的努力。它源于一种感觉：恰如其分，增一分则多，减一分则少。技术中的美不一定非得需要原创性，在技术中，无论是形式，还是短语的选用，都大量借鉴了其他语言。如此看来，具有讽刺意味的是，设计所依赖的往往是对"陈词滥调"的组合和操纵。尽管如此，一个美的设计总是包含一些意想不到的组合，并以其适切性震撼人心。

技术就像写作、演讲或高级烹饪，存在不同程度的流畅性、表达能力和自我表现。一个建筑界的新手，就像一个学习外语的新手一样，即使有时不太适合，也会一遍又一遍地一直使用同样的组合，或者说同样的短语。一个熟练的建筑师深谙域的艺术，他会摒弃纯粹的语法规则，而诉诸直觉知识，使它们组合、搭配在一起。一位真正的大师会挑战极限，他会在域中赋诗，会在其所使用的惯用组合中留下自己的"签名"。

技术大师实际上非常难得，因为技术的语法不像语言的语法，它变化迅速。技术语法最初是原始的，只能被模糊地感知；当组成它们的基础知识发展时，它们得以深化；当发现了新的匹配良好的组合，或者发现了设计在日常使用中遇到的困难时，它们就进化了。这个过程永无终结。结果就是，即便是此中的专家，也不能完全跟上它们在"域"中组合的原则。

在一个域中投入巨大的成本会让设计师较少考虑从所有可用的域中选择组合。保罗·克利（Paul Klee）认为，艺术家会自觉去适应自己调色盘上所能配出的东西，"画家……不会让画去迎合世界。他自己会去迎合画作"。[11]技术也是这样，设计师总是从他们了解的域中着手建构技术。

参与的世界

如前文所述，一个域或者一个技术体提供一种表达的语言、一套内容和实践的词汇表，设计师可以从中作出选择。计算技术（或数字技术）是一个集合，是一个巨大的词汇表，包括硬件、软件、传输网络、协议、语言、超大规模集成电路、算法，以及所有和它们相关的组件和实践。所以，我们可以把计算技术，或者任何与之相关的域看作是一个仓库，它们随时准备服务于某种特殊用途。

我们可以将这个仓库假设为装有元素或功能的工具箱。但是我更愿意设想它是一个王国，那是一个可以在其中建功立业的世界，一个为了能在其中完成某个任务而呈现的世界。

一个域便是一个想象的王国。在那里，设计者可以在思维中想象自己能做什么。那是一个充满可能性的域世界。电子设计师知道他们可以扩大信号、转换频率、减少噪音、调节载波信息、设置定时回路，还可以利用许许多多其他可靠的操作。他们依据电子世界实现内容的可能性去思考。此外，如果他们是专家，他们应该非常熟悉这个世界，因此他们几乎可以自动地组合操作并预见结果。那个发明了微波激射器（激光器前身）的物理学家查尔斯·汤斯（Charles Townes），曾经在波和原子共振相关的操作和设备上花费了许多年时间来研究以下内容，包括：场下离子分离、共鸣室、敏感高频接收器和检测器、微波光谱。然后他将这些功能应用到他的发明中去。专家们沉浸在域的世界里，就如我们写信的时候沉浸在文字中一样，他们的精神沉浸其中。他们着眼于目的性，然后在头脑中进行每一步操作，这很像一个作曲家构思出一个主题，但是却要回过头来诉诸乐器去表达它。

在另一个意义上，一个域也是一个世界。在那里，有设计者和用户都可以接触并完成的真实的任务，在那里，常规的操作是可能的，使用过程也总是相同的。一些目标（行动或者业务流程）也以实体形态进入世界当中。想象一下，图像处理专家通过扫描使图像进入“数码世界”，或者是用数码技术摄制图像。一旦有对象被从一个操作传递到另一个，就会继续下去，完成转换，有时还会与这个领域内其他的活动和对象结合起来。在数码世界里，影像变成了数值、数据，因此它可以接受数字化操作，包括颜色校正、锐化、去饱和度、变形成广角效果、增加背景等。当操作结束后，对象再次以这种加工过的形式，被应用到现实世界中。作为被操作过的数据，加工过的影像被翻译成真实世界的视觉图像，又一次展现在计算机屏幕上，被储存起来或被打印出来。

无论是域还是它的世界，这种操作过程构成了域的真正的有用性。

我们可以认为某些东西“下到”了一个具体环境，在那里成为对象，被进行各种操作，然后再令它回到“上面或外面”并被使用。现场交响乐可以通过麦克风设备被带到“下面”的电子世界进行加工（电子化操作，例如电子记录），然后再被带回到“上面或外面”，重新进入物理世界，以声音的形式被“演奏”出来。

对于域世界来说，每个域里的所能完成的事情是不同的。某些域世界提供了特别丰富的可能性。数码世界可以对任何简化成数字符号的东西进行操作，不管是建筑设计还是摄影图片，或是飞机的操控系统。它提供了一个巨大的运算操作或运算次序，以及逻辑步骤的系统。而这些操作的速度可以非常迅速，数字电路中的变换是会以超快的变换速率进行的。

另一些域世界则很有限，在能力方面会有所限制，但是它们可以更有效。在18世纪后期，运河提供了一个域世界，在那里用载货驳船更便宜，也更容易运输大宗商品，如煤炭、粮食、木材、石灰石，甚至牲畜等。这样，大宗货物的运输就通过进入运河世界完成了。准确地说，这意味着货物已经离开了道路和陆地域，而进入了驳船、船夫、船闸和纤道构成的水的世界。当然，在那里，东西的移动是缓慢的。但尽管如此，由于它的移动是流动性的，因此与陆地比较，这种移动更容易。在这个域中的货物可以流向不同方向，向后或横着停泊在水中突起的陆地上。一些货物被卸下来，腾出空间，再装上另外一些货物，到达目的地后，运输还可以继续进入公路或马车运输的域，而那里曾是运输业开始的地方。

已经成为历史的“运河域世界”实际上功能很单一，它其实只提供一种功能：运送物料。甚至这个功能也只能在开挖运河的地方才能实现。但是它在成本–收益方面是成功的。在河道运输之前，陆地运输只能依赖笨拙的牛车蹒跚在颠簸的土路上，而19世纪早期河道运输方式在英格兰推广以后，煤的运输成本就下降了85%。

一个域的世界里最容易完成的事情体现出了这个域的力量。只要你能将某事还原为数字式描述，从而进行数学操作，计算就可以行之有效地进行下面的工作。只要你能将货物装载到驳船上，并用它来运送货物，运河世界就能开始运作。电子学（非数字类）的有效性取决于运动的显现要依赖于电子的运动。域们各善其事。当然，从原理上讲，你可以在不同的域世界中完成同样的任务，但是功效可能会不一样。应用数字域可以很容易列出顾客名单，但是如果你非要用电子域去完成它，你

可能需要用不同的电压代表不同的字母，然后按照感应电压和输出的振幅顺序来安排电路，这当然可行，但很笨拙。你甚至可以用运河域世界来分列顾客名单：比如每艘驳船可以作为一位顾客的标记，然后驳船被向前拖动就如同字母缓慢显现。这也是可行的，但是它肯定不是运河世界最有效的应用方向。

如我所说，不同的域世界提供各自擅长的、互不相同的潜能。每个世界提供各自最容易完成的一套操作。所以这是很自然的，一个对象或业务活动，要进入一个以上的世界，各尽其能地完成整体工作。光学数据传输提供了一个光子的世界，通过光纤网络发送信息。这里的主力是光子（或光量子）包。这些信息很容易被携带，并且以接近真空中的光速迅速移动。但是这里有个问题：光子不像电子那样可以携带电荷，因而它们是中性的，不易操作，而且由于部分光缆中的光总是被吸收，因此信息每隔数英里就会衰减。所以这些信息必须不断“重复”或放大，才能得以继续传输，这意味着光子流必须被适当地还原和调整。

在早期，光子世界并没有成为传递信息的直接手段。信息经常退出光子世界，转而进入电子世界，因为信息在那里更容易被重新调整、放大和切换。但是电子世界太慢了，它必须依赖电子的移动，而那需要对电磁场作出反应，因而不能瞬时完成。这就像每隔几英里，信息就要被带下高速公路，到一个便道进行电子操作，然后再把它重新放回高速公路一样。系统本身是可行的，但是持续的离开和再进入光子世界提高了成本而且降低了速度。直到适用的光子放大器（即掺铒光纤放大器，EDFA）在20世纪80年代后期诞生后，这个问题才得到解决。

常规经验是，当任何一项活动离开一个域世界并进入另一个域世界的时候，其成本就会累积增加。在海上运输货运集装箱并不昂贵，但将货物从铁路运输域转移到船上使货运进入“集装箱世界”，这就需要铁路端点、码头、集装箱装卸起重机和装卸业务等烦琐而昂贵的技术。这种“过渡技术”通常是一个域中最棘手的部分。过渡技术会产生延迟和瓶颈，并提高成本，但它们又是不可或缺的，因为它们不仅使域有效，而且控制什么可以进入或离开它的世界。我们可以这样看待域：它包含着一些核心操作，比如蒸汽驱动的、便宜的海运集装箱运输，还包括处于域的边缘，围绕这些中心操作的缓慢的、笨拙的技术。与其相对应的域世界允许其在活动开始时进入，在活动结束后再离开，比如说甲板和龙门吊。总体上讲，它们是昂贵的。

我此前曾说过，域反映着它们创造的那个世界的力量，但是它们也

同时反映着它的局限。将设计域定在数字技术中的建筑师可以让设计思路的变化瞬间呈现出来，并且可以同时将材料成本计算出来。但是数字域会有它自身微妙的偏见，只有真实世界可量化的部分能够进入数字世界，并在那里获得成功。因此，数字建筑设计可以轻而易举地产生拱形，或者以流畅的数学曲线呈现倾斜，但是正如建筑设计评论家保罗·葛贝格（Paul Goldberger）所说，数字化设计对“魅力、闲适、留白都缺乏耐心”。[12]数字世界可以计量表面形状，却无法计量时尚。或者我应该说，时尚目前还不能被量化。如果它能够被量化了，即如果你可以通过显示器上的移动滑块控制你想要的魅力程度，那么现有的域就需要扩展了——不过到目前为止，这还是不可能的。一个世界的“不能”，便是这个世界的局限所在了。

我将在第8章重新谈论域或者说复数的技术，去探究它们是怎样产生以及怎样随着时间发展的。现在我们只需认识到，当我们将技术理论化，我们必须看到技术的中间层（技术体）是在不同于单个技术的规则之下运行的。这些技术体或者域，决定着某个特定的时代的可能性，它们引发一个时代的标志性工业，它们是工程师可以有所成就的世界。

在这个世界中，没有什么是静止的，待完成的东西随着域的演进及其基本现象边界的扩展而不断变化着。这暗示着创新不是发明以及对其的应用（例如，计算机、运河、DNA或者芯片的发明和应用），而是在新的可能世界中，将旧任务（例如，会计、运输或者医疗诊断）不断地进行重新表达或者再域定的过程。

THE NATURE OF TECHNOLOGY

第5章 工程和对应的解决方案

几乎所有的设计都是某个已知技术的新版本，只有在必要的情况下，才需要一项全新的设计。工程师在寻找解决方案的过程中，把适宜的构件选择出来，让它们组合在一起工作。设计就是选择，组合只不过是工程的副产品。工程的解决方案，又成为发展新技术的新构件。

到目前为止，我们探究了技术的本质，它的工作原理以及它最深层的含义等问题。在此基础上，我们形成了一套关于技术的逻辑和框架，用以说明技术在世界中是如何被建构和操作的。在本书的后几个章节，我将继续应用这个逻辑去探索技术是如何形成并进化的。

我们现在已经不再将技术看作是作为整体的一块铁板了，而将其看成是具有内部解剖结构的事物。事实上，一旦我们接受技术是建构的产物，即由零部件或组件组合而成，我们就不得不这样看待它们了。这样一个内部视角能否使我们从与以往不同的角度去看待技术这个问题呢？

我认为会，这种不同至少会体现在两个方面。第一是关于技术在其生命周期当中如何进行自我修正。如果我们从技术外部将技术看作是一个整体的对象，那么个体技术，如计算机、基因测序、蒸汽机似乎是相对固定的。它们的生命周期可能表现为从一个版本到下一个版本的变化。如计算机从阿塔纳索夫-贝瑞计算机到埃克特和莫齐利的电子数字积分计算机（Eckert and Mauchly's ENIAC），再到电子数据计算机（EDVAC），这些技术可能会以这种方式非连续性地进行改变。但是当我们从内部来看技术的时候，我们会看到某一技术的内部组件一直都在变化，比如替换零部件、改进材料、改变建构方法、对技术所基于的现象有了更好的理解，或者随着母域（parent domain）的发展，有了新的元素可以利用等。所以，技术并非是总体上很静态、只是偶尔发生变化的事物。正好相反，技术是一种非常易变的东西，它是动态的、活的，会随时间发展而不断进行构成和发生变化。

第二个不同之处是我们如何看待技术的可能性（我现在是在“集合”的意义上谈技术）。从外部看技术，每项技术看起来都是在完成某个目标：如果想测量，我们有测量的方法；如果想导航，我们有全球定位系统。我们可以应用测量方法完成某个具体任务，应用GPS完成另外一个具体的任务。但是这么看技术，在“理解技术是什么”时是有局限的。因为技术不仅是为了提供某种特定功能而存在，它实际上还提供了一个组合或编程的词汇表，这个词汇表的存在使技术可以提供无穷无尽的新颖方法，去实现无穷无尽的新颖目的。

如果这样看技术的话，结果就会有很大的不同。比如可以设想一下未来某天，计算机技术已经消失了，考古学家们挖掘出了一个20世纪80年代的破旧的苹果电脑，于是他们急忙回到实验室，插上电源，这个珍

贵的盒子就闪烁着开启了，那么他们应该立刻就能发现几种能够执行的功能：用于文字处理的MacWrite程序，用于影像制作的MacPaint程序，或者一个旧版的电子制表软件。这里的每个功能都是可用的，可以很好地执行特定的任务，因而调查者可以在很长时间内使用这个机器去执行每个任务。

但是实际上，这个苹果机还可以提供更多用途。在苹果机的内部，蕴藏着它的工具箱，这是一套内部的命令和函数，它们可以使一般的目的被程序化执行。这些命令可以以某种方式被组合起来，从而创建以前没有的新命令和函数（功能）。这些新命令自身可以被命名，并且被用作未来组合的新组件。只要从故纸堆里对苹果机的相关知识进行充分研究，研究者们就可能学会如何进入其内部命令系统。他们会搞清楚如何提取命令，并把它们重新组合，去完成新的任务，或者将它们作为新的程序命令的组件。在这个时刻，会出现一个质的飞跃，研究者们会发现他们可以操控苹果机了，他们可以用一小部分基础命令进行编辑，这些基础命令再以新的方式进行新的无限的组合，从而简单命令可以被建构为许多复杂命令。

这时的机器就不再只是那个提供几种独立功能的机器了，它现在可以提供的是一种语言表达。到此，一个可能的新世界已经被开启了。

技术通过改变内部组件的结构进行适应，以及通过新组合产生新结构，这两个主题将是本书接下来不断重复的主题。但是别忘了，我们的主题是：技术进化是如何通过组合现存技术来产生新技术，以及如何通过现存技术去驯服可能形成新技术的现象的。我想要探究的是：这个过程到底是怎样发生的，对此我会在第9章更加详细地加以论述。

现在让我们沿着这个思路继续下去，我们需要留意技术写作中的两个反复出现的问题。一个是达尔文机制应用于技术进化中的程度，即在什么程度上，技术的新“物种”会从某种旧技术中通过变异而产生出来并且被选择？另一个是，在什么程度上，托马斯·库恩的观点会被应用到技术中？库恩认为，业已公认的科学范式会随时间的推移变得更加完善，直到遭遇所谓的“异常”，进而发生更新换代的情形。那么这个观点也适用于技术吗？

我们还需留意的议题是：创新。“创新”是技术中另一个棘手的词。通常在实践或尝试新点子时，如果出现了一些进步或提高，不管它多么微小，创新都有可能会被唤醒。熊彼特使用这个词（对我来说，有些费解）来说明发明被商业化的过程，而我则会在它的一般含义上，也就

是“新颖性”这个意义上使用这个词。这种新颖性会表现为几种形式：给定技术中的新的解决方式、新技术自身、新的技术体，或者在技术集合中加入新元素等。在后面的章节中，因为“创新”这个概念太分散、太模糊、不利于更有效地说明问题，所以我选择去探究新颖性或创新的每个具体类型。

一个好问题可能是成功的开始，为此我们需要一个关键的、足可以引领我们本章和下一章探讨的问题来开始进一步的探究。进化的发生源于新技术的不断形成，它们通过将已有技术作为组件来形成（forming）新面貌的方式，表明它们是脱胎于此的。那么，这种“形成”到底是怎样发生的？新技术的产生机制是什么？一个显而易见的答案是，它是通过某种根本性创新（radical innovation）过程（如果你愿意，你也可以称之为发明过程）来实现的，简单来讲，确实可以这样认为。但是别忘了，新组件的元素也来自于日常性的标准工程。根本性创新与日常标准工程放在一起，初看起来会令人感到有些讶异，因此我想挖掘一下这个过程究竟是如何发生的。

首先让我澄清一下什么是我所说的“日常标准工程”。

标准工程

工程师们的日常工作到底是什么？总的来说，他们设计并且建造人工物。他们也开发方法、建立测试设备并进行研究，以找出材料被运用以及解决方案在实践中得以履行的方法。他们在专业的研究所或实验室中深化对操作对象的理解；他们调查失败的原因、寻找修复的对策；他们管理法律事务并提供建议，对顾问委员会提供咨询与服务；他们对相关问题，也即那些好似需要反复讨论、思考及担忧的问题进行审慎地斟酌。

> 技术思想前沿
>
> 标准工程是执行一个新项目时，在已知可接受的原则下聚集方法和设备的过程，是对已有技术的新的计划、试制和集成过程。

工程师们为所有这些活动奔忙。但是我想集中讨论的是标准工程[1]，即执行一个新项目时，在已知可接受的原则下聚集方法和设备的过程。有时这个过程被称作“设计和建造（construction）”，有时则被称

为“设计和制造（manufacture）”。但无论怎样称呼，它们都是对已有技术的新的计划、试制和集成过程。它不是指斜拉桥或者飞行器的“发明”，而是指设计和建造一个新版的斜拉桥，比如日本的多多罗大桥或者新版的空中客车。方便起见，我将这种活动简称为“设计”。

几乎所有的设计项目都是在规划和建造某个已知技术的新版本。这就像几乎所有的科学活动实际上都是将已知的概念和方法应用于新问题一样。当然，这并不意味着标准工程是简单的。我们可以按照困难程度列出一个谱系：从应用惯例和标准组件的传统项目，到那些需要实践或实验的项目，再到那些拥有真正难以突破的界限并需要应对一些特殊挑战的项目。在后面的例子中，我会着重讨论这个谱系中更具挑战性的一端。

标准工程或者说设计项目，到底应该包括什么？其基本任务是需要找到一个形式（form），或者说一套已建构好的程序集（architected assemblies）来实现目的。这意味着要用一些可用的概念框架和目的进行匹配，然后再进行现实的集成。这是一个过程，而且经常是一个冗长的过程。教科书里通常会讲到三个阶段：先从一个总体概念出发，然后细化出可以完成这个概念的集成件，最后实行制造或建造（这个过程中会伴随一些必要的反馈）。这里我们可以再次借用递归性来描述标准工程这种沿层级演进的过程，即从总体概念层次到单个集成件，再到次级集成件，再到它们各自的零部件，接下来每一个部分的构成也是上述过程的重复性进行。

事情大致来讲就是这样，但仅仅是大致来讲。设计过程在沿着层级向下演进的同时，也会从需求特点或者需求物（desiderata）向外演进。目的本身决定总体概念的样式，总体概念又决定着需要什么样的核心集成件，核心集成件决定需要何种次级集成件来支撑它们，次级集成件又决定它们所需的组件。20世纪60年代末，波音747的总工程师约瑟夫·萨特（Joseph Sutter）的一段描述正可以说明这个过程[2]：我们想要设计一种搭载350名乘客的飞机。我们设想建造一个宽阔的单甲板。由于在载客区会有9～10人并肩而坐，因此机身的长度也就被大致确定了。我们试图优化机翼使其达到我们需要的吊装能力、飞行里程以及燃油效率。而机翼跨度首先要满足空气动力学的要求，还要达到初始巡航高度的要求，以及保持合理的降落速度以便飞行员有足够的时间着陆。

注意一下这里的顺序，需求作为飞机的主要目的开始显现——要搭载350名乘客，然后向外扩展构成集成件，每个集成件的需求决定了下

一步要做什么，每个层次的集成件又需要互相匹配和相互支撑。

可以设想，这个过程有时可以进行得如同预先计划好的一样完美。有些项目确实如此。但是在难度谱系中具有挑战性的那一端，就不一定那样利落清晰了。通常整体概念不一定是一个，有可能存在好几个整体概念，而且有一些概念还是直到实验阶段或者甚至细节设计阶段才发现是不可行的。即使一个概念被选中了，它也必须被转化成集成件或者工作组件，而其中许多组件都需要特别设计。设计者不可能总是事先就预测出它们的实际性能。先前的版本也可能表现出某些未预料到的小毛病：可能不如期望的那样好，可能根本不运作，或者可能耗用了更多的质量、能量和成本。为了获取更好的解决方案或者原材料，必须对其进行修改。在一个集成过程中，不可预知的低效率必须通过调整其他部分得以提高。可以说，一项设计就是一系列折中的过程。

要使过程进行下去，除了需要大量的支持者之外，还要对理念、集成件和单元组件进行测试和平衡，并克服不断被揭示出来的困难。[3]只要有一个核心集成件不好用，就有可能需要重新开始整个计划。如果项目特别复杂（例如，登月计划）则可能需要分步骤进行不同版本的实验，每种技术都分别建构在之前技术的基础之上。

一个项目一旦进入到一个未知领域，许多小问题的出现就变得不可避免了。比如，1965年的波音747还处于构思阶段，那时，它的巨大的重量要求一个比以往任何时候都要强大的动力系统来支撑。这不仅要求更大的涡轮风扇，而且要求更高的涵道比（从风扇吹出的风与从发动机流出的燃气之比要从以往的1：1达到几乎6：1）。为此，普拉特·惠特尼集团打算重新设计JT9D引擎。一个巨大的涡轮风扇被安装在压缩机前面，其直径为244厘米，能提供77%的推力。新的动力装置使性能实现了一个飞跃式的进步，但是它的创新性特征又引发了另外一系列困难。它的变容量定子（用来帮助控制通过压缩机叶片的气流）本来是由可移动的链接控制的，但是这些改进会产生周期性的卡停（解决的办法是自由使用WD-40），而要使这个装置产生更高的压缩温度，则又需要一个更好的涡轮机冷却手段。

最糟糕的问题是将发动机安装在机翼上选择何种方式。当发动机高速运转时，它会被向前推动并产生“弯曲”，引起外壳产生轻微的椭圆化变形。“根本原因在于引擎太大、太重，导致它起飞时会弯曲。”JT9D项目副经理罗伯特·罗萨蒂（Robert Rosati）解释道。发动机高速运转造成的挠度并不大，大约是1毫米，但这却足以引起高压压缩机叶片的底部

与外壳之间的摩擦。这类问题并不危险，一定量的摩擦也是可以接受的。不可接受的是由此带来的效率和可靠性上的损失。普拉特和惠特尼尝试了几种修正的方法，如加固外壳、使用抗磨损椭圆形封条，但都没有成功。最终的解决方案是安装一个反向的Y型锚固件，这在本质上是一种转移推力、减少弯曲的手段。挠度因此减小到80%，足以应付上述问题，但是这个小挫折使整个波音747的研发进程受到了拖延与阻碍。对于这样顶尖级的项目，类似的挫折并不常见。

在开发波音747的年代，实现新的设计需要手动完成庞大的计算量，手绘大量的明细图以及手工制作许多实体模型。现在，计算机接替了这些工作。计算机可以在瞬间将设计理念转换成明细图和部件要求，可以创造虚拟模型，有时甚至可以指导零部件的制造。但是即使有计算机辅助，这个过程也必须要有人的参与。这是因为过程进行中需要决策，而决策是不能依赖机器的。设计者必须对概念、结构、材料、强度、比例和容量的适合程度作出评价和判断。而像这样在一个项目的许多层级都要做恰当匹配就一定需要人的协调。

但是对于规模较大的项目来说，协调通常是很困难的。大规模项目的组件可能是由不同团队甚至不同公司设计的，这样一来，他们之间就需要进行平衡。一个团队的解决方案对于另一个团队来说可能反而是障碍，因此需要充分讨论来加以协调。从这个角度看，标准工程成了一种社会组织形式，一切都变得不那么清晰理性了。历史学家托马斯·休斯（Thomas Hughes）强调，一项新项目成功与否，它是否能形成可见的设计物，很大程度上依赖于围绕其周围的利益关系网：工程团队、融资系统、投资者以及其他参与者，他们从项目中获得或失去的包括权力、安全、威望等东西。因此，设计与发展是与人高度相关的组织和行动过程。

解决问题的工程

此前我说过，工程师花费了大量的时间，有时候几乎是他们全部的时间去解决问题。为什么非得如此呢？教师教学、法官断案，那么工程师为什么不专司工程？为什么他们要花费那么多的时间去解决问题？

在我们看来，运气不好可能是一种解释，例如波音747遭遇到的那个不可预见的挫折。无法满足不同利益集团的利益也是一个原因。技术挫折或人的原因当然很重要，但是它们并不是最主要的，工程之所以和

解决问题紧密相连，有着更系统的原因。

正如我所定义的，标准工程与已知技术有关。这使得每个设计实际上都是已知技术的新版本或新案例（JT9D就是喷气式发动机的一个新版本）。但是一个新案例或一项新设计只有在技术的某个方面需要变得不同的时候才会被需要（如果不进行新设计的话，设计中已完成的部分将被迫放弃，建构过程也将停止）。新的设计可能是需要达到一个新的功效水平（例如，JT9D）；或者是需要一个不同的物理环境；或者有了更好的零件和材料的选择；还有可能是市场发生了变化，因而需要某个技术的新版本。总之，无论哪种情形，一个新设计只有在必须的情况下才会被实施。

这意味着一个新项目总会抛出一些新问题。完整的回应（即完成的项目）通常就是一种解决方案，也就是需要在具体的理念指导下对集成件进行恰当地组合以完成这个给定的任务。我们可以说，一个完整的设计就是对特定工程问题的特定解决。

想要全面解决问题，就一定会遇上许多新情况，因为每个层级上的集成件的选择都需要重新考虑，并重新作出相互和谐的设计。某些集成件或模块当然可以单独进行修改，但是总的来说，当一个现存技术正在建构成一个新的版本的时候，每个层级，每个层级的每个模块都需要重新加以考虑。如果它不能和其他部分或预期的希望相匹配的话，必然要被重新设计。而这里的每一项设计又都会抛出它们自己的问题。所以我们可以更准确地说，一个完整的设计是针对一系列问题的一系列解决。

这么说并不意味着每个解决结果都能令人满意。不良的解决结果可能引起持续不断的问题。波音737就曾因方向舵失灵（专业术语叫“异常”）问题而备受困扰，这个问题曾引起过一次坠毁。工程师们用了很长时间去理解和解决这个问题。设计问题在技术的一个生命周期中常常不能够被彻底地解决，从莱特兄弟的飞行器到现代F-35喷气式战斗机，为飞机提供合适的控制系统一直都是一个难题。由计算机控制的自动驾驶现在成为一种解决办法，但是这种解决方法本身还处于完善当中，需要和应用它的新机型共同进步。

组合与解决方案

所有这些都解释了为什么在难度谱系更具挑战性的那端，工程的主要任务是解决问题，以及为什么工程师会不断地遇到问题。

但是这里又有了一个新问题。如果工程是关于问题的，那么是什么构成了一个或一套针对问题的解决方案呢？我曾经说过，一个解决方案就是能完成一项给定任务的恰适组合。工程中的任何创造都是一次为达成特定目标的建构，这是一个将众多元素组合起来的过程。所以我们可以这样重新发问：一个解决方案是如何被建构的？这种建构又是如何与组合关联起来的？

实际上，对我们来说，这才是议题的核心。这本书通篇思考的是作为组合的技术——假使设计也能成为一个组合的过程，组合是如何在设计中发挥作用的呢？当然，工程师会选择适宜的组件并把它们放置在一起，他们组合它们，让它们共同工作。但是这并不意味着他们一开始就明确地知道如何组装所有的东西，知道他们将组合出什么。工程师只是简单地将自己投身于完成某个目标、满足某些技术条件或人，以及解决那些连带问题的过程之中。其中，精神劳动的部分包括了选择，即选定哪些部分共同组合成一个组件。组合不是工程创造过程的目的，而是选择的结果，是为了产生技术的一个新实例而完成要素聚集的结果。这样看来，组合实际上是工程的一项副产品。

这类似于思想表达的过程。现代心理学和哲学都告诉我们，思维在一开始都不会产生于语言之中。我们从无意识层面拔升出我们的理念——思想，然后用词语的组合去表达它们。思维存在了，它的语言表达随后才会到来。

如果会说几种语言，你就能明白，或者应该说能感受到这点。假设你的公司正在莫斯科做生意，和你一起谈判的人都只会说俄语，或者只会说英语。你想要说点什么，于是你用俄语说了。片刻后，你又用英语表达了同样的想法。这个“想法”某种程度上是独立于你如何将它说出来的。你有了一种想说什么的意向，然后经由某种无意识的过程找到话语进行表达，其结果是你最终呈现出什么样的说话方式：它可能简短而自然，比如在一场对话中；也可能显得冗长，类似你在准备一个演讲时需要将内容一点一点拼接起来。无论哪种，它都是一个与某个目的相联系的想法或概念的组合，然后需要用句子、短语，或只用基本的单词进行表达。你在创造这种组合时并非刻意为之，但实际上你确实在“组合”。

这个过程和技术是一样的。设计师先对某事产生意向，为了表达这个意向而去挑选一个工具箱或语言，为了将这个存在于“心灵的眼睛”（mind's eye）中的意向呈现出来而去预想出概念和功能[4]，然后找出合适的组件去组合，从而最终达成那个想法。预想常常会同时产生，

或者也可能会在修正之后出现。我们会在第6章更详细地探讨这样的创造过程是如何进行的，但是现在，我们要关注的是它和语言的关系，先有意向，然后是完成的方法——组件的恰当组合。所以，设计即表达。

这暗示着工程及其寓所是一个创造性的领域。工程通常被认为比那些更重视设计的领域，如建筑或音乐，更缺少创造性。当然，可以说工程是日常性的，但是我们也可以说建筑同样是日常性的。因为从原则上看，工程中的设计，与建筑、时尚或者音乐中的设计，以及所有我们能想到的创造性的方式其实并没有什么不同，它们都是某种组合、某种表达。

与其他创造性领域相比，工程创造性被低估的一个原因是，与建筑和音乐不同，公众没有被训练如何去欣赏一个具体的、制作精良的技术。计算机科学家霍尔（C. A. R. Hoare）在1960年创造了快速排序法[5]，这是一种真正优美的创造，但是他不能到卡耐基大厅去表演快速排序从而赢得欢呼和掌声。还有另外一个原因，技术工作大部分都隐藏起来了，它们倾向于深藏不露，罩在某个罩子下面，藏在某个程序里，或者躲在某个工业流程中。谁能看见一位手机设计师是如何解决一个具体问题的？总之，这些工作对外行来说完全是不可见的。

创造偶尔还是可以被看见的。20世纪的前10年，瑞士工程师罗伯特·马亚尔（Robert Maillart）创造了一系列桥[6]，作为通用技术，它们根本不新鲜，但是却和勒·柯布西耶（Le Corbusier）或者密斯·凡·德·罗（Mies van der Rohe）的创造一样具有创新性。在桥梁被用繁复的雕饰和厚重的石墙建成的年代，马亚尔的大桥显得优雅轻盈。直至今天，他的作品也显得非常现代。土木工程师大卫·比林顿（David Billington）在描述罗伯特·马亚尔于1933年设计的施万德巴赫大桥时，曾说它是"有史以来最美丽的两三座混凝土大桥之一"。它看起来与其说是要跨越沟壑还不如说是漂浮在那里，它那样纤细，却有着无可比拟的创新性。

作为建筑，施万德巴赫大桥并没有采用什么新形式：它依然是由固定在拱上的竖向构件来支撑桥面。这是马亚尔时代广为接受的结构形式，它根本没用什么新材料，也没用什么新的结构构件——钢筋混凝土在19世纪90年代中期就已经被应用了。马亚尔完成他如此优雅的表达时，所用的却几乎都是最普通的技术手段。经过大量的几何学分析之后——尽管他不是天才的数学家，他明白了需要大大加强桥面的刚度，这有利于把重力荷载（假如桥的一端有一辆卡车）均匀地分散在大桥的结构上。设想一个桥的实体模型，其支撑拱是一条金属带，其两端被牢牢

地固定在桥墩上，但是中间部分是弯曲的。现在将桥面，一个平板，置于拱的顶部，两端也固定在桥墩上，由固定在拱上的竖向构件来支撑。在桥面的一端施加荷载会使这端向下，而另一端则会翘起。如果桥面是软的，这个重力荷载就只能作用在施力端，而这是他们不希望看到的。但如果桥面是坚硬的，桥面的硬度将使反作用力作用在未加载端，从而产生一个向上的推力，这样一来，荷载将被更均匀地分散在整个结构上。

正是这个“解决方案”使得桥拱和桥面的重量可以很轻，而同时也有足够的强度——而且其中又蕴藏着优雅。而重量轻也使得建桥时可以使用很迷你的脚手架。马亚尔还学会了熟练地运用新型的钢筋混凝土，从而把自己从厚重的砖石建筑中解放出来。他摒弃了几乎所有的装饰，这也是他的建筑直到现在看起来还很现代的一个重要原因。这种形式既有效又经济，而且具有创新性。但是它成功的原因却远不止这些。零部件和材料组合在一起，在整体上呈现了一种流动与和谐。这里完成的当然是一个技术产品，但毋宁说，它更是一件艺术品。

我并不想将标准工程及其主要的实践者们浪漫化。你在马亚尔的大桥中所发现的技巧并不来自于“天才”，而是来自于知识和技能的经年累月的积累，这正是马亚尔拥有并为之着迷的东西。并不是所有的标准技术的例子都和马亚尔的一样。大多数的项目依然坚持为标准问题提供标准的解决方案。需要的尺寸或规格可能会有一些变化，但是重新计算和设计不外乎还是依据一些必要的标准模板。即使最具日常性的项目也是针对一个问题或一组问题的一个或一组解决方案，同时一定也具有创造性。

这里有了一个结论。设计就是关于解决方案的选择。因此，设计与选择有关。如果一项技术的所有部分都被诸如重量、绩效和成本严格限定的话，选择看起来就可能很严格。但是限定经常使得问题的解决更加复杂，继而需要调集更多的部件来完成。关于一个复杂表达的选择的工作量，即解决方案和解决方案的解决方案（次级解决方案），它的数量是巨大的。技术的任何新版本都可能成为后来大量不同建构方式的潜在要素。

实践中的建构方式比理论上的要少，因为工程师倾向于重复使用以前曾经用过的解决方案——短语和表达，并且他们倾向于使用可获得的现成的组件。所以，由同一从业者完成的新项目通常少有新奇的解决方案。但是由许多不同设计者同时参与的项目则会产生许多新颖的解决方

式，它们会出现在诸如实现目标的概念设定方面，在域的选择上，在组件的组合过程中，在材料、建构方式以及生产技法方面。所有这些创新聚集起来，推动现存技术及其领域的进步。不同解决方案或次级解决方案的经验以这种方式被牢固地聚集在一起，促使技术随着时间而变化和进步，其结果就是创新。

经济史学家内森·罗森伯格在谈到微小进步的聚集现象时说："改进是通过不为人注意的设计和工程活动而获得的，但是它们构成了巨变的基本内容，并且在经济生活中为消费者带来了福利。"[7]从这个意义上说，标准工程对创新贡献良多。

标准工程就是一个认识过程。

未来的技术构件

标准工程还有其他功能。它为技术的进化作出了贡献。读者根据我前文所述的内容可能已经预想到这一点了。大多数时候，工程问题的解决方案都是具体的，而且不是作为一个整体加入到技术的指令系统中的。但是一些解决方案会被偶然地多次重复使用，从而使其本身变成了目标，继而在未来的技术建构中成为新的元素。

如果你看过任何一本工程手册，你就会发现许多标准问题的解决方案。比如《机械和机械设备图集》[8]就给出了"耦合旋转轴的19种方法"，以及"15种不同的凸轮机构"。另外一个电子类的手册，则图解了5种振荡电器：阿姆斯特朗振荡器、科耳皮兹振荡器、克拉普振荡器、哈脱来振荡器、瓦卡振荡器。这类手册提供了标准解决方案来解决那些重复性的问题，也可以为特定的用途进行修改。有时候这些解决方案是以适当的发明的形式出现的，最后成了解决未解问题的正式答案。但是它们更多时候是因实践者找到了新的方式，即找到了一种新的、巧妙的组合现存组件或方法的方式，来解决一个标准问题。如果设计结果特别有用，就会被其他技术采用。通常它会先在共同体内部进行推广，之后再成为通用模式，这时它就变成了一个新的技术构件。

这个过程很像理查德·道金斯所讨论的模因（meme，即文化基因）[9]的运作方式。模因，正如道金斯最初设想的那样，是类似信念、流行语、时装等文化表达的单元，它们在社会中被模仿、重复和传播。成功的解决方案和理念就是以这种方式在工程中起作用的。它们在从业者中也是这样被模仿、重复和传播的。它们是组合的后备元素。

> 技术思想前沿
>
> 一个解决方案如果被使用的次数足够多，它就成了一个模块，并因作为适用于标准用途的模块而具有包容性，它自己也成了一项技术。

事实上，如果被使用的次数足够多，一个解决方案，一个成功的组合就成了一个模块。它会获得自己的名分，并因作为适用于标准用途的模块而具有包容性，同时它自己也成为一项技术。当一个新的术语可以概述某些复杂的想法的集合，并且成为词汇的一个新的组成部分的时候，就会有与之并行的相应的语言产生出来。“水门”或者“慕尼黑”就是源于对一系列特定的、复杂的政府过失或谈判过程的概述而建构起来的。现在，“某某门”和“慕尼黑”已经固化为独立的名词，用来表示政府过失或政治绥靖。它们已经成为语言中的构成模块，被加入到构成英语的整体元素中了。

这种机械主义的解决方案能产生建构达尔文主义的模块吗？正如我所描述的，它听起来很达尔文主义。工程问题的解决方案是多样的，其中较好的会被选择并自我传播。但是我们需要注意：解决方案并不是像生物进化那样一步一步地变化而来的，它们是快速聚集起来的组合，并且以解决问题为目的。

准确来讲，解决工程问题的过程产生了新的解决方案——某种新颖的组合，但按照达尔文的说法，它不是渐变的，而是以突变方式到来的。较好的方案会被选择，并依照达尔文方式通过工程实践进行传播。最后，一些解决方案将成为建构新技术的元素。产生技术构件的最基本的机制是组合。达尔文机制在随后的选择过程中起作用，其结果是只有某些解决方案能够存活下去。

顺便说一句，这种选择并不意味着技术中最好或最适合的解决方案总会存活下来。当针对工程中给定问题的几种解决方式出现的时候，我们可以认为这种竞争是为了应用——为了能被工程设计者所采用。随着解决方案越来越得到普及，它也越来越显眼，从而更可能被其他的设计者所采用和改进。小的偶然事件，例如谁在什么时间和谁提到过某个解决方案，谁的方法在贸易杂志上被提到过，谁推销过某个解决方案等，都有可能推动某个解决方案，使其领先一步。因此它会被其他设计者更进一步的采用，进而在它的域的实践中被“锁定”（lock in）。在过程中获得主导地位的解决方案肯定是有其优点的，但是不一定非得是竞争中

最好的。它可能是非常偶然地在竞争中占据优势的。

一波流行引发另一波流行，然后被锁定，这是一个偶然的过程。这一点我已经在前面集中描述过了，所以不再赘述。可以肯定地说，流行的技术（或解决方案）更趋向于获得进一步的优势，继而被锁定，所以在技术“选择”中，存在一个正反馈过程。

你可以在核电站的核反应堆设计案例中看到这种正反馈。这里的主要问题之一是冷却材料的选择：用什么将热量从反应堆的核心部分传输到涡轮机？另一个问题是慢化剂（moderator）的选择：用什么控制中子的能量水平？在核电技术发展的早期，对此曾经有很多建议——“啤酒”，一位工程师说道，这是目前唯一还没试过的慢化剂。有3种方案曾经被广泛地开发：轻水（H_2O），同时用于冷却和慢化；重水（D_2O），同时用于冷却和慢化；用气体（通常是氦气或者二氧化碳）进行冷却，用石墨进行慢化。

不同国家、不同公司都经历过这些方案开发过程。加拿大倾向于使用重水方案，因为它有水力发电那一套技术用来进行重水制取；英国曾试验过气体-石墨方案，但都没有成功。美国则同时使用了几种方案。最后，美国海军在海曼·里奇欧文（Hyman Rickover）上将的领导下，开发了核潜水项目。[10]尽管钠可能更有效，并且体积也更小，但是钠在水中会爆炸，暴露在空气中则会燃烧。由于担忧潜艇内的钠泄漏，里奇欧文最终为他的潜水艇选择了轻水冷却。此外，工程师更熟悉压缩水，因而没有对液态钠系统进行更多尝试。

1949年，苏联爆炸了第一颗原子弹。美国的一个反应就是想通过他们拥有的核反应堆来彰显其核优势，所以美国原子能委员会在里奇欧文的建议下，在宾夕法尼亚的海港将本来打算用于航空母舰的反应堆重新设计成供陆地使用。新的反应堆和以前用于海上的反应堆一样，也采用轻水慢化。后来，西屋电气和通用电气开发出来的反应堆也都是采用轻水模式。史学家马克·赫兹加德（Mark Hertsgaard）对此评论道：“这为轻水模式奠定了里程碑式的开端，其他方案无法与之匹敌，从而使该产业将它的商业前景实际上置于某些专家认为的经济和技术上都较弱的一种设计之上。”

到了1986年，世界范围内（苏联除外）的101个反应堆中，有81个都建立在轻水基础上。轻水解决方案占据了主导地位。[11]它虽然始于小的偶然事件，但将继续占据未来设计的主导地位。最终流行起来的技术是被从几种可能的“解决方案”中选择出来的。但是，正如后来的研究指

出的那样，它不一定是最好的。

在我研究“标准工程”之前，我不曾奢望它能对技术创新和进化有多么大的贡献，但是正如你从本章看到的，我现在改变看法了。标准工程中的每一个项目都会使一系列问题显现出来，每一个解决的结果都是一套对应的解决方案。可用的解决方案被建构并在实践者中传播，其中一些可能变成技术名词，进而变成未来技术的建构元素或模块。标准工程对创新和进化都贡献良多。

在第6章，我将转到另一个话题：新技术（不是指已有技术的翻版）是如何诞生的。换句话说，就是发明是如何形成的。在这之前我需要澄清一点，许多目的系统（贸易公约、侵权行为法、工会、货币系统）都是全新的，它们没有所谓的“前身”，但是它们也不总是被刻意发明出来的，如此说来，它们存在于标准工程和发明之间。那么我们需要如何对待这样一个范畴呢？

我们注意到，上述目的系统是一个缓慢的随着时间呈现的过程，这和标准工程中解决方案的产生很相似。工会实际上不是被“发明”出来的，它们是从中世纪的雇佣工互助联合会，即共谋兄弟会的形式中衍生出来的。从早期的记录中你甚至可以目睹它们的形成过程：

> 剪羊毛的人“常常在城里所有同业市场之间转悠，他们会借此进行一些密谋，比如命令所有同业中人停止工作，或者不为自己的主人服务，直到主人和奴仆之间或市场之间达成某种共识”。[12]

我们看到了工会的一种新形式的形成过程（其内核则来自14世纪）。几百年来，始于社会实践的工会经历着成长、夯实、发展，并且应环境的要求呈现出不同的形式。

这类偶然形成的事情并不罕见，在此无须赘述。考虑到我们的主张，我们只需注意到，新的目的性系统（作为解决经济或社会问题的实践或者惯例）可以在不经意中产生，而且如果确实能起作用，它们将继续成为更广泛的系统的组件。在本书中，我们的主要精力还是放在严格意义上的主动创造新技术，即发明这件事情上，所以让我们继续追问它们是如何形成的吧。

THE NATURE OF TECHNOLOGY

第6章 技术的起源

新技术可以是根据某个目的或需要发现一个可以实现的原理，也可以从某一新现象出发，找到如何使用这种现象的办法。原理可以借用，也可以是先前概念的组合，或者由理论而来。只有将概念转化为现实，一项新技术才真正诞生。科学和数学中的原创，与技术没什么两样，因为它们同属“目的性系统”。

达尔文生物进化论最需要回答的核心问题是："新物种是如何产生的？"[1]类似地，我们理论的核心问题是：根本性的新技术是如何产生的？当然我说过，不能简单地用达尔文进化论来解释技术。比如，我们不能说喷气式发动机是由此前的发动机经过自然选择获得的微小变化的积累而产生的，也不是简单地将现有的技术碎片在头脑中或现实中囫囵地加以组合就能产生的。正如熊彼特所说："无论你如何重组邮政马车，你永远不能因此而得到铁路。"[2]这并不是否认组合原理，而是强调了创新会涉及更多的秩序性，而不仅仅是纯粹的随机偶然性的结果。

那么，新技术到底是如何产生的呢？

我们实际上也可以这么问：发明是如何发生的？奇怪的是，尽管这个问题很重要，但是在关于技术的现代思考中却还没有令人满意的答案。最后一个将发明理论化的尝试发生在20世纪30年代，但是接下来的数十年间，这个主题却被搁置了。其原因在很大程度上是由于"创造行为"的核心部分被认为是没办法进行评估的。因此，发明在技术中占据的位置就像是心理学中的"心理"或"意识"，人们愿意谈论它，但是并不想真正解释它是什么，教科书常提到它，但也仅是一带而过，以避免对诸如"它是如何运作的"这样的问题作出解释。

我们确实知道一些关于新技术是如何诞生的故事。我们知道（大多数是从社会学研究中获得的）新技术是由社会需求形塑而成的；它们主要来自于标准域以外的经验；它们更经常地在鼓励冒险精神的环境中产生；它们更容易在伴随知识交换的过程中产生；它们经常会在网络中得到促进。没错，这些解释都对，但是它们在解释"一项新技术是怎样形成的"这个问题上，除了指出"种子发芽是由于有合适的土壤"之外，并没有给出更多的洞见。

因此，对于技术产生的最核心部分，在历经几十年才能形成的经济结构的核心部分，以及我们所成就的人类福祉的基础部分，却依然是一团谜雾。

构成经济的装置、方法和产品到底来自于哪里？那到底是个什么样的过程？这是我们的问题。

什么样的技术才算新技术

首先我们需要弄清楚，什么样的技术能被看成是一项发明呢？什么样的技术才符合根本性新技术（那些在深层意义上与以往不同的技术）的定义呢？我在这里将根本性的新技术定义为：针对现有目的而采用一个新的或不同的原理来实现的技术。原理就是做某件事情的操作方法，使某事运作的基本方式。

> 技术思想前沿
>
> 新技术是针对现有目的而采用一个新的或不同的原理来实现的技术。新技术是在概念当中或实际形态当中，将特定的需求与可开发的现象链接起来的过程。

对于我们通常认为的发明，这样的界定对吗？20世纪70年代，计算机打印是通过行式打印机来实现的，其实质就是带有固定字母的电子打字机。随着激光打印机的发现，计算机打印开始通过引导激光在硒鼓上“打印”文本来实现了。在20世纪20年代，飞机动力是通过活塞式螺旋桨推动的，随着喷气式发动机的发明，飞机动力改由汽油涡轮发动机产生反推力来实现了。这采用的是不同的原理。20世纪40年代，数字计算是通过机电方式实现的，随着计算机时代的到来，数字计算通过电子中继电路得以完成。在这些案例中，新技术的形成（激光打印机、涡轮喷气发动机、计算机）都是源于一个新的或不同的基本原理。

原理中的某个变化使发明，即根本性的新技术产生的过程，从标准工程中分离出来。如此一来，我们可以分辨其中关键的差异仅仅是改进还是真正的原创。我们说，波音747是波音707的改进，而不是一项发明。它是对一项现存技术的发展而不是全盘应用新的原理。我们说，瓦特蒸汽机是纽可门蒸汽机的改进，因为它只是提供了分离式冷凝器这个新的组件，但是没有应用新原理。（当然，在商业上，有时候改进比纯原始性发明更重要。）这样看来，判断一个技术是否是真正的发明，我们仅需要通过判定对当前的目的起作用的是否是一个新的、不同以往的根本性原理即可。而这样一来，却可能导致出现一个判定的灰色地带，马亚尔的加筋甲板（stiffened deck）是一个改进的组件，还是一个新的原理？它两者兼有。依据根本原理新颖性的程度可以产生一个连续统一体，而它存在于标准工程和根本的新颖性之间。

我们依然没有找到一个关于新技术如何形成的准确理论，但是我们现在有了一个有效的标准来衡量什么算是合格的“新”，我们在此基础上

继续深入探讨。

一项新技术是怎样产生的呢？

我们现在思考的基点是新技术（发明）应该应用新原理。我说过，原理就是应用某种现象、概念或理念。所以当我们说技术建构在新原理之上，实际上是说它建构在新的或不同的一个或几个现象之上。这强烈地暗示了新技术是从何而来的。新技术是在概念或实际形态当中，将特定的需求与可开发的现象链接起来的过程。我们可以说，发明是将需求和一些现象链接起来，并能令人满意地满足那个需求的过程。（当然与标准工程比起来，这个原理或这个现象的应用对于那个目的来说一定是新的。）

我发现勾勒出这个链条很有用。链条的一端是“需求”或“目的”；另一端是能达成需求或目的的基本“现象”。在两个端点之间是一套完整的解决方案，即新的原理或现象被用来实现目的的过程。但是如何使新原理恰当地起作用，也是个颇具挑战性的过程，这需要它们继续寻找各自的解决手段。整个过程通常是以系统或集成的方式使问题的解决成为可能的过程，我们可以将其想象为解决问题的链条上的一系列环节。

这个比喻还远没结束。每一个环节反过来又都有其自身的任务，并可能因此需要接受属于它的挑战。它可能因此又需要它自己的次级链接，或者次次级解决方案。是的，不用感到惊讶，链接过程也是递归性的。它包含链接-解决-进一步的次链接-进一步的解决，并且它们可能又需要它们自己的解决方案或者发明。我们可以将发明看作集成这些链条的一个过程。这个过程会持续下去，直到每个问题和次问题都能找到现实的解决方式，直到链条完整为止。

在实际中，链接过程会相当不同。一些发明可能是单兵作战，另一些则是团队行动；一些拥有巨资资助，另一些则依靠微小资金的努力；一些需要经过经年累月的试错，并显然产生了一系列不能完全实现目标的中间版本，另一些则好似无中生有般突然呈现。

技术思想前沿

发明有两大模式：

- 肇始于链条的一端，源于一个给定的目的或需求，然后发现一个可以实现的原理。

- 发轫于链条的另一端，从一个现象或效应开始，然后逐步嵌入一些如何使用它的原理。

但是无论发明过程多么变化多端，最终我们都可以将它们归为两大模式。它可能肇始于链条的一端，源于一个给定的目的或需求，然后发现一个可以实现的原理。或者，它也可以发轫于链条的另一端，从一个通常是新发现的现象或效应开始，然后逐步嵌入一些如何使用它的原理。无论哪种模式，其过程都要等到将原理转化成工作元件之后才算完成。

两种模式会在很大程度上相互重叠，所以没有必要将两种过程都详细地加以描述。我将主要探究始于可见需求的那一类模式的过程，然后我再简短地补充一下始于现象的另一种模式。

找到一个基本原理

让我们假设发明起始于一个目的，为此要找到一个实现某种需求的解决办法。这个需求可能来自于某种经济机会，一种关乎潜在市场利润的可能性，或者来自经济环境的变化、社会挑战、军事需求。

需求通常并不来自于外部刺激，而是源于技术自身。20世纪20年代，飞机设计者们认识到他们可以使飞机在高纬度稀薄的空气中获得更高的速度。但是在这样高的纬度上，往复式发动机甚至是压缩空气超动力发动机都无法得到足够的氧气，而导致螺旋桨缺少必要的“咬力”（bite）。为了解决这个问题，就需要一个不同于活塞-螺旋桨（piston-propeller）的原理。

这种需求产生的方式很具典型意义。它们出现时常常只有极少数实践者知晓，同时可行的解决方案也还未出现。假设这时已经有了解决方案，那么标准技术就足以应付。但现在不行，因而如何定义问题本身就变成了挑战。那些接受挑战的人会遇到类似需要满足某个需求或需要克服某种限制的情况。他们很快会将挑战简化为需求——一个需要解决的技术问题。喷气式发动机的原创者（originators，由于“发明者[inventor]”这个词常常蕴含“古怪而孤独地进行工作的人”之意，所以我不打算使用它）弗兰克·惠特尔和汉斯·冯·奥海因都意识到此前的活塞-螺旋桨原理的局限，因而需要寻找不同的原理。但是他们将这种需要重新表达为一个技术性的问题——一个需要满足的要求。惠特尔为此选择去

寻找某种动力设备，要轻巧且高效，可以适应高纬度的稀薄空气，如果有可能，甚至可以不要螺旋桨。冯·奥海因则去寻找“稳定的奇热动力学流动过程”[3]，意思是“当空气进到这个系统时，在遇到对马赫数敏感（mach-number-sensitive）的发动机元件之前会减速”。至此，需求已经变成一个被详尽描述的问题。

现在需要继续去寻找合适的解决方案。思想者（我暂且将发起者设定为一个人，而通常的情况都是几个人同时工作）开始专注在问题上，开始调查、搜索能够通过进一步发展来满足必要条件的可能性。这种搜索通常是概念性的、宽泛的、执着的。牛顿的一段很有名的关于他的引力轨道理论的评论是：顿悟来自于“连续不断的思考”。这种连续的思考允许无意识工作，有可能从过往的经历中回想起某个现象或概念，而且它提供了一个无意识的警觉，当一个可供选择的原理或界定问题的不同方式出现时，叩门声就会轻轻地响起。

这个阶段所追求的并不是完整的设计，正如我前面说过的，这时需要寻找一个基本的概念，即关于何种效应或效应组合作用时正好会解决问题的理念，以及关于达成这个理念所需要的工具手段的设想。

每个经过认真思考得出的备选原理又会带来自己独有的困难，这些困难将抛出新的次一级的概念性问题。因此“新障碍”等于是对待解问题进行了缩小或再定义。一旦思想者认识到一小部分问题可以被解决，其他大部分的解决方案就会跟进，或者至少会比较容易地落实到位。这是个在不同层级之间来回跳跃的过程，一会儿需要测试原理在某个层级上的可行性，一会儿需要解决它们在较低层级上产生的问题。

上文提到的整个过程如同预先策划登山线路一样。抵达顶峰相当于解决问题，预想一个基本原则如同设定一个从某个确定地点出发的比较可靠的整体（或至少大部分）的攀登线路。山上是障碍重重，充满了未知的冰瀑、裂谷、峭壁、雪崩和坠石。每一个新的原理（或者说攀爬的总体规划）都会遇到需要克服的困难。

此时，递归性又一次上演了，因为每一个障碍都变成了自身的次级问题，并需要各自的解决方案（次级原理或者次级技术）。直到完成了从出发地到山顶的可行的攀登规划之后，才意味着获得了整体的解决方案。

当然，我们可能曾经领略过这座山的某处风景，也就是说，某个次级技术是我们知道的，那么其解决方案就可能被采用。如此一来，整个

过程更像将已知部分连缀在一起，而不是作为先锋去寻找一条完全陌生之路。这个过程也是递归性的，整个事件内部是级联关系，它形成了一个前进的计划，或者对我们来说，是一个技术的预想过程。

这些候选路径，这些原理，又是从何而来的呢？

它们来自几个方面。有时候原理是借用的，它们本来满足的是别的目的，或使用的是另外的域。惠特尔在1928年曾考虑了几种可能性[4]：火箭推进力、转动式喷嘴的反作用力、使用螺旋的涡轮推力（涡轮螺旋桨飞机）、活塞引擎推动的管道排风机，以及所有这些问题引发的次级问题。这里的每一种可能性都借用了本来服务于其他技术目的的原理。

> 技术思想前沿
>
> 原理从何而来？有时候原理是借用的；有时候原理来自于以前概念的组合；有时候原理来自于对过去的回顾；有时候原理是和现存功能性结合在一起出现的。原理来自于已有的其他设备、方法、理论或功能之中，它们从来都不是无中生有的。

有时候，新原理来自于以前概念的组合。1940年，英国战争期间需要找到传输雷达微波的有效方式。物理学家约翰·兰德尔（John Randall）和亨瑞·布特（Henry Boot）一下子想到磁电管的原理[5]——一种被用来服务雷达目的而产生微波，利用磁场控制电流的圆柱状电子管。为此他们将电磁管的高能量输出和电子速调管利用共振腔扩大微波的优点组合了起来。

有时候，原理来自于对过去的回顾，或者从同事的谈论中偶然获得，或者由理论而来。实际上，兰德尔曾经偶然在书店看到了海因里希·赫兹的《电波》（*Electric Waves*）的英译本。这本书使他想到了圆柱谐振腔——基本上就是赫兹在他的书中所分析的三维的线圈共鸣箱。

有时候，原理（一种概念上的解决）是和现存的功能性结合在一起出现的，每一次出现解决一个次生问题。1929年，欧内斯特·劳伦斯（Ernest Lawrence）找到了使带电粒子加速以实现高能粒子对撞技术的方法，即粒子可以被电场加速。但问题是，当时人们还不知道如何获得产生高强度电场的极高电压。直到有天晚上，劳伦斯在大学图书馆浏览学术期刊时，发现了挪威工程师罗尔夫·威德罗（Rolf Wideroe）写的一篇文章。罗尔夫·威德罗的想法是用低电压交流电使粒子反复振荡进行加

速，这样就回避了高电压问题。他建议通过一系列首尾相连的管子来传输粒子，管子和管子之间有小的缝隙。管子的安排要恰到好处，即粒子一定要在交流电的峰值通过缝隙的时候同时到达。但是这意味着当粒子运动越快，管子的长度就需要越长。劳伦斯看到了这个方案的精妙之处，但是他算了一下，如果想达到他要的能量，管子就要伸到实验室的窗户外面去了（按照现在的观点，管子的长度需要达到3千米）。威德罗的想法在劳伦斯看来不具可行性。

但是就像当时任何一位物理学家一样，劳伦斯知道，磁场可以引起带电粒子在回路中运动。我问我自己："是否可能用两个电极管一刻不停地输送阳离子（粒子），通过某种合适的磁场排布，让它们在电极管中来回运动。"[6]换句话说，可以通过只使用两根管子，将它们弯成两个半圆，中间有缝隙，然后使用磁场驱使粒子在这个环形回路中来回运动。接着他让威德罗的恰到好处的交流电通过缝隙，这样粒子就在每次通过缝隙时被加速了。当它们不停地旋转运动时，就可以被加速、盘旋，最后被高能释放。

这个原理最终演变成了回旋加速器。它的实现过程是：将问题从如何获得高电压转变为如何使用威德罗提出的低电压交流的次级原理，然后再利用劳伦斯的方法，用磁场来大大降低空间需求的次次级原理。此处获得的原理是在已存在的碎片（现存的功能）的基础上建构而产生的。

在所有这些案例中，原理都是来自于已有的其他设备、方法、理论或功能之中，它们从来都不是无中生有的。[7]在发明的创造核心，呈现的是"挪用"（appropriation）的特征，是某种半意识形态的精神借鉴。

有时候，原理来得很迅速，只需花费很少的心力。但是更多时候，整个问题都藏在思想背后，被一些困难所困扰，并没有现成的原理摆在眼前，而且这种状况可以持续数月甚至数年。

解决方案有时候会突然出现，就像查尔斯·汤恩斯（Charles Townes）在讲述他发明量子放大器的过程时说的那样："灵感倏然而至"。惠特尔也曾写道：

> 我在惠特灵时，突然间我的脑海中出现了这样一个念头，去找一个替换活塞发动机的涡轮机（用来驱动的压缩机）。这种变化意味着压缩机要有一个比我对活塞发动机的期望高得多的压力比。简言之，我又回到了燃气轮机上，但这一次是推进

喷气式飞机，而不是推进压缩机。当想法初现时看起来怪怪的，我花了很长时间才找到那个概念，而突然间它显得那样明显而且简单。计算结果使我确信这远远超出我的预期。[8]

灵感的显现就如同一个疏通堵塞的过程，经常是一下子就通畅了。要么是总原理找到了可以与之匹配的次级原理，要么是一个次级原理为主原理的应用扫清了障碍。这是一个连通的时刻。它在问题与能够解决问题的原理之间完成了链接。

使人感到惊奇的是，对于发现者来说，这个洞见是如此完整，让人觉得在无意识中各部分已经被组合完好了一样。而且它一来，大家就“知道”它是对的——那是一种对其解决问题所具备的正当性、优雅性、非凡的简洁性的觉察。因为它总是从一个人的无意识中涌现出来，所以洞察力通常来自个体，而不是团队。它一般不会在活动中或狂想时出现，却常常在静寂中到来。

灵感力的到来并不是过程的结尾，而只是一个标志。概念依然必须被转化为可行的技术原型。就像作曲家虽然在头脑中已有了主题，但是依然需要通过演奏来将之一起表达出来一样，原创者必须将工作组件组装起来，才能完成他的概念的表达。

概念的物化

概念物化的进程很可能早已开启了。一些设备或方法的概念通常在经验中已经被建构起来，某些基本概念也可能早已经在实践中被试验过了。所以发明过程的第二步和第一步通常是有重叠的。将概念完全实现，意味着详尽的细节建构；关键组件必须被制造出来、作出平衡并进行建构；要选择恰当的测量工具；要进行理论计算。这些建构需要背后的鼓励和资金支持。竞争在这个阶段是有帮助的。事实上，如果各竞争团队觊觎的是同一个原理，那么接下来的竞争一定非常激烈。

将理念变为现实的过程会带来许多挑战，这些挑战或许曾经在头脑中被很多次地预想过，而如今必须在真实世界中面对它们了：在提出解决方案的过程中，有时会经历失败，如零部件可能不适用，也许需要重新进行设计，也许必须要进行实验等。发明的第二步主要是寻找次级问题（subproblems）的解决方案，其中会包含许多标准工程的特征。

这一步所面对的挑战可能是巨大的。20世纪60年代晚期，施乐

（Xerox）公司的盖瑞·斯塔克伟泽（Gary Starkweather）曾经探索一种不会受制于缓慢的逐行打印机（本质上的一种大型打字机）的限制，而能够直接打印数字化字节（例如，由计算机制造的图像或文本）的方法。最初，他曾经想到过用激光将图像“打印”到硒鼓上这一核心理念。但是在将概念付诸实际操作时，他遇到了几个困难，其中有两个关键性的难题亟须得到根本的解决：为了使打印过程可以商业化，必须将扫描文稿的时间控制在几秒钟之内，也就是说，为了得到高分辨率的扫描结果，需要激光束能够以每秒50万次的速率通过开关的通断得到调制，从而使黑点或白点留在硒鼓鼓面上。但在当时，以这样的速率调整激光是不可能的。同时，对于这个任务来说，激光和镜头上的所有模块（module）都太重了，这将带来过大的惯性，从而导致在以每秒数千次速率的扫描过程中出现持续反复的物理振荡。要完成既定的技术目标，这两个问题都必须得到解决。

斯塔克伟泽通过应用由一个压电单元驱动的偏振滤波器，开发出一个非常快的快门装置，从而解决了开关高速通断的问题。[9]通过利用一个旋转多楞面镜，可以使激光束移动，而激光模块却可以保持不动，这样就解决了惯性问题。当多棱镜转动的时候，每一面都可以扫描硒鼓上的一窄行，这很像灯塔发出的光束横扫过地面时的情形。但是这个解决方法随即带来了它自己的次生问题。斯塔克伟泽经计算发现，棱镜相邻的镜面的紧密度公差（tight tolerance）必须是准确的6弧秒，否则相邻的扫描行就不能很好地相互弥补，从而出现扭曲。但是，加工到那种精细的程度在成本上是不可行的。一款精心设计的柱透镜（斯塔克伟泽是主修光学的）解决了这个问题，它可以确保相邻行不至非常紧密，这样即使镜面轻微错开也不要紧。

当你听完斯塔克伟泽的故事后，印象最深的应该是他所面临的选择。每个次生问题都可能有几种解决方案。斯塔克伟泽选择解决方案、检验它们的可行性，并努力从中组成和谐的整体。当需要找到次生或次次生问题的源头时，他就向下拓展这一梯式递归进路，如果这些问题得到解决或被放弃了，他就再向上回溯。这个过程几乎总是很漫长，只有当获得了必要的知识，并且次生技术带来的挑战已被成功地克服后，才能向前迈一大步。前进的方向总是沿着运行恰切的版本向前延伸的。

最早的导航装置是一个很具说明性的例子。尽管它最初看起来并不完善，然而它诞生的那一刻却弥足珍贵。所有关于其起源的叙述都会记下那粗糙的集成闪烁出生命火花的时刻。在这一刻，原理得到了自证。

成功了，一个里程碑过去了，让我们为那些时刻而欢呼。“1954年4月初的一天，我和我的学生们正在进行一场研讨会，吉姆·戈登（Jim Gordon）突然闯进来”，汤恩斯（在谈论微波激射器的发明时）回忆道，“他翘课是为了完成一个实验。他高呼他成功了！我们立刻停止了讨论，和他一起奔向实验室去看振荡的证据，然后就开始庆祝他的成功了。”[10]

最初的呈现可能真的很微弱，但是随着进一步的努力、自组织性的修复以及随后呈现的更棒的建构，巩固的工作状态就浮现出来。这时，一个新的基本原理就进入到一种基本可靠的状态了，即它具备了物理形态。所有这些都需要时间——它考验着赞助者和主管们的耐心。在这段时间里，最为必要的人类品质就是意志力，一种要赋予原理以生命，使其成为有效实体的意志力。到目前为止，一个新装置或方法又成为进一步开发并进行商业应用的候选者了，如果足够幸运，它就有可能作为某种创新进入经济领域。

发明，作为一个过程，现在完成了。

基于现象的发明

我们简单地谈两句发明的另一个过程——始于现象的发明过程。这个过程也是现象和目的的链接过程，不一样的地方在于，过程开始于现象一端。通常，当人们注意到了一个现象或拥有了关于该现象的理论时，就意味着有了一个如何使用它的理念、一个原理。和始于需求动机的过程一样，接下来的过程也同样是建构出支撑组件来将原理转化成现实的技术。

始于现象的发明过程似乎应该简单些了。一个现象可能直接就暗示了一个可应用的原理，而且通常的确如此。但是有时这种暗示也是不清晰的，比如1928年亚历山大·弗莱明（Alexander Fleming）那个著名的发现就是如此。他当时注意到一个现象，即一种霉菌（后来被证明是青霉菌的孢子）中的某种物质可以抑制葡萄球菌的生长，他随即意识到这可以用来治疗感染。这一存在于现象和应用之间的联系在回顾时很清晰，但实际上还有许多人，比如物理学家约翰·廷德尔（John Tyndall）在1876年，安德烈·格拉提（Andre Gratia）在20世纪20年代都曾经先于弗莱明注意到了这个反应，但他们却都没有预见到它的医疗用途。弗莱明“看到”这个原理是因为他在第一次世界大战中曾经是医生，对于战地

感染造成的伤亡曾感到震惊，因此他更容易发现一个看似无用而实际上意义重大的现象。

即使应用原理清晰可见，技术实践的转化工作也并非轻而易举。通常的情况是，如果效应是新颖的，就不易被理解，同时相应的技术工具可能还未被开发。例如，将青霉菌现象转化成可用的治疗方法首先意味着对青霉菌中的活性成分进行隔离和纯化；接着需要弄清楚它的化学结构；然后需要经过临床试验检验其疗效；最后还要进行生产方式的开发。这一整套步骤完成下来，已经远远超出了弗莱明的能力范围，它需要集结高度专业化的生物化学专家们来共同完成。最终这一计划是由牛津邓恩病理学院的霍华德·弗洛里（Howard Florey）和恩斯特·柴恩（Ernst Chain）领导的生化学家团队来执行的。13年之后，弗莱明的发现才转化成实际可行的技术产物——盘尼西林[11]。

我曾经说过，发明或者始于需求，或者始于现象，但是读者可能会反对说，有许多发明并不是这样的。莱特兄弟确实不是从需求或现象出发的。自力推动飞行器的愿望以及实现它的两个基本原理（通过轻体内燃机推动、通过固定式机翼飞行）在许多年前就存在了。事实上，这种情况并不少见。基本技术原理和对应的需求经常一起被发现，而能将其变为现实的次生原理和元器件则可能要等到数十年后才会被人们发现。莱特兄弟所要做的是解决阻碍原理被转化成可行技术的4个关键的次生问题。通过认真地实验和无数次尝试，他们终于解决了飞机控制及其稳定性的问题，发现了机翼对飞机上升的举足轻重的作用，建造了轻体推动系统，开发出了高效推进器。他们1903年的那次动力飞行不能作为对一次“发明”过程的完整解释，它只不过是前人踩踏出的漫长小路上的一个标志而已。

莱特兄弟的案例并没有形成其他不同的发明模式，它只是我描述的两种模式的一个变种。基本原理有时会自然呈现，有时会突然出现。困难之处在于如何使原理正确地发挥作用，这可能需要漫长的努力。

什么是发明的核心

关于发明我已经说得够多了。发明的核心在于发现合适的可行性解决方案，即“看见”合适的工作原理。剩下的，夸张点讲，就是标准工程了。有时候原理显而易见并容易借鉴，它会自然而然地呈现。但大多数情况下，它需要进行深思熟虑地心理联想，那好似一个在头脑中进行的

链接过程。

那么这种心理链接过程到底是怎样发生的呢？[12]

回到劳伦斯案例。我们注意到，开始时他并没有想到后来那个最终解决方案，即将一个电磁铁与一个产生于在两个D型容器之间的振荡射频电场组合起来。他思考的是如何将可采取的行动和可利用的效应（功能）组合起来，并达成解决方案。换句话说，他是将问题和解决方式联系起来，并且想象当某些组合构成后会发生什么。

劳伦斯的洞察力固然是深邃的，但这只是因为他所利用的不是我们特别熟悉的功能。从原理上看，劳伦斯的问题和我们日常面对的问题其实并没什么不同。当我的车还在修理店里而我却想去上班时，我可能会想：我可以先坐地铁，之后换乘出租车，或者我可以搭朋友的车，又或者我可以在家工作，只要我能在我的小窝里腾出块地方就行。我在我的日常功能储备库中搜寻，选出一些加以组合，并考虑每种解决方案所带来的次生问题。当我们看到这种方式应用在我们日常的生活当中时，会感到这种推理并不神秘。其实在发明中，它也没什么不同。可能它发生的领域是我们不熟悉的，但是那一定是发起者所熟悉的。发明的核心是心理联想。

我在谈论劳伦斯那类的心理联想时用到了功能性。发起者进入功能库中去想象，如果某些功能组合起来会发生什么。但有时联想直接来自于原理自身。原理经常跨界工作。例如，哪里有波（声波、洋流、地震波、无线电波、光波、X射线、粒子），哪里就有干涉（两个或两个以上的波能产生叠加）、频谱、共鸣系统（以其固有频率振荡）、折射（当进入一个新介质时，波会改变方向），以及多普勒效应（如果波源相对于我们运动，频率会有可被感知的变化）。所有这些都会提供一个可以利用的概念——原理，它们反过来也可以从传统“域”中借用过来，并应用到新“域”当中。发起者想到需求功能：如何才能测量运动？如何才能在一个特定的频率中制造一个稳定的振荡？然后他会从某些已知的领域开始联想，并借用其中的原理。兰德尔借用了赫兹的感应线圈原理，并且设想它在三维模式中如同圆柱谐振腔一样工作。当发起者需要一个特定功能的时候，他们可以回过头联想他们知道的、在某一领域内产生过相应功能的原理。这一机制的核心称为原理转换（principle transfer），它是“发现”的一种类比，这是另一种心理联想的形式。

我说发明的核心是心理联想，但并没打算排除想象。恰恰相反，发起者必须具备足够的想象力去理解问题，这是第一位重要的，然后才是

预见它如何被解决、有多少种解决方式、必要的组分及结构是怎样的，以及如何解决随机出现的那些还不可见的次生问题。但是这类想象绝不神秘。发起者的共同之处既不是“天赋”，也不是某种特殊的能力。实际上，我不相信有天赋这类东西，我只相信对巨大的功能和原理库的占有能力。发起者通常很熟悉他们即将应用的原理、现象的相关实践或理论内容，惠特尔的父亲是一位机械师和发明家，因而惠特尔从小就对涡轮机很熟悉。

然而，发起者不仅要掌握某种功能性，并在他们伟大的创造中一劳永逸地用一次，还要更频繁地面对冗长的功能集结过程，以及由此带来的无休止的、如同钢琴五指练习般的针对微小问题的试验过程。在和功能性打交道的这个阶段，经常可以看出发起者试图利用某种现象的蛛丝马迹。在查理·汤恩斯的伟大发现完成的5年前，他就曾在他的备忘录中提到微波无线电“已经可以利用如此短的波长，以至于已经和对分子共振有丰富研究的量子力学理论和光谱分析技术有了部分重叠，这可能会对无线电工程有巨大帮助”。而他正是用分子共振来发明微波激射器的。[13]

我们可以看到功能性的专业知识聚集过程，通常这个过程在发起者眼中都是理所当然的。生化学家凯利·穆利斯（Kary Mullis）认为他的聚合酶链反应计划（从单一DNA样品复制出含有大量DNA拷贝的DNA链）非常简单。“它太简单了……所涉及的每个步骤都已经完成了。”[14]但是穆利斯的“简单”解决方案是“通过往复促成与带有特殊DNA顺序的独立各股相杂交的两条引物的交互延伸来放大DNA”。这就意味着首先要找到一段DNA原本，然后标出待复制片段的首尾，接着将DNA双螺旋分离成两条单链。一旦原本被加入，双链就可以使用一种被称为聚合酶的酶开始复制，连接互补成分形成两个新的双螺旋。重复这个过程使得新的双螺旋以2、4、8、16的倍数无限地进行复制。穆利斯口中的“简单”，只是对像穆利斯这样的对DNA功能非常熟悉，操作也非常有经验的人来说，才是简单的。

因果性金字塔

我在本章已经描述了作为微观过程的发明，即一个人（或几个人）提出一种新的做事方法。但是它一定是发生在某个情境中的，换句话说，新技术一定衍生于此前已经存在的组分或功能之上。这样的观察可以使我们如同透过广角镜一样看到更全面的发明过程：新技术实际上是

由先前作出铺垫的一系列设备、发明和理解的堆积而形成的山峰。

事实上，我们可以假定任何新设备或新方法都是一座通向顶峰的因果性金字塔。那是一座应用共同原理的技术金字塔，一座包含所有对此新技术有所贡献的先驱技术的金字塔，一座包含那些使新技术成为可能的支撑原理或组分的金字塔，一座包含使这些新技术成为可能的、但是曾经一度是新现象的金字塔，一座包含用于新技术中的仪器、技术以及制造过程的金字塔，一座包括先前工艺及其理解的金字塔，一座包含用来描述现象的语法以及所用原理的金字塔，一座包含在这些层级上人们之间互动的金字塔。

在这座因果性金字塔中，特别重要的是随时间增长的知识的积累，既包括科学形态也包括技术形态的知识。正如历史学家乔尔·莫基尔和埃德温·雷顿（Edwin Layton）指出的，这种知识不仅存在于工程实践自身当中，而且也存在于工科院校、学术团体、国家科学院、工程院以及公开发行的出版物当中。知识构成了新技术呈现过程中至关重要的基础部分。

这一更为宽泛的视角[15]与我们前面的论述并不矛盾。这个因果性金字塔对发明的宏观过程的支撑，非常像战争中后勤系统对部队的支撑。事实上，可以用历史因果性的解释来替换我这种个人用法。这就像解释滑铁卢战役时，其军事作战兵力、文化、培训和设备，其先前的成就，以及他们的补给线都可能成为解释战争胜利的原因，但我们往往将赢得战争的原因更聚焦于在实际战斗中发生军事冲突的紧急时刻所作的决策和采取的行动等方面。

说新技术拥有因果性历史并不表示它们的出现是可以预先确定的。发明取决于奇思怪想和发现新现象的时机，还取决于新需求的出现，以及对此作出回应的人的洞察力。同时，由于所有发明都会受到因果性金字塔支撑，这也意味着当必要性和需求的碎片都一一铺垫到位之时，一项发明就将显露。

这种时机呈现过程中粗略的“就绪状态”使得一项新技术很少拥有唯一的发起者。几位颇具代表性的发明家可能会几乎同时想到同样一个原理并准备付诸实施。众多的努力以及关键部件的加入使得我们实际上很难从“第一”这个意义上来讨论“发明”。更多时候，我们可能看到部分原理以与从前相关的方式出现或者只是以往形式[16]的重新体现，我们无法很好地领会它们，但实际上它们有共同的出身。我们总能看到一个系列产品的原初版本是由不同的工人互相借鉴着开发出来的，只是随着新的

次级技术的出现，设备和方法才从最初的粗糙状态逐渐变得精致起来。以计算机为例，它实际上并不很像“发明”。我们可以说克劳德·香农（Claude Shannon）“看见”了利用电子传递线路进行算术运算的基本原理，然后原理的实践转化版本被建立起来，之后通过互相借鉴，加上组件的连续改进，最终完成了计算机的发明。因此，计算机的发明不是一蹴而就的。

如同上述情形展现的那样，设定一项所谓的“发明”是困难的。现代关于技术的有些论述已经谈到这一点。计算机先锋迈克尔·威廉姆斯（Michael Williams）就认为：

> 与人类发明相关的活动中根本就不存在什么“第一”。如果你加上足够的形容词来描述，你总是可以声称你找到了你认为的第一。例如电子数字积分计算机（即ENIAC）经常宣称自己是“第一台电子、通用、规模化、数字计算机”。但是，在你有一个正确的陈述之前，你一定要加上所有这些形容词。如果离开它们中的任何一个，那么，像阿塔纳索夫-贝瑞计算机，“巨人”（colossus，美国海军水下声呐探测潜艇警报系统的代号），楚泽（Zuse）的Z3计算机（德国康拉德·楚泽开发的第一台采用二进制和布尔逻辑的可编程计算机），以及许多其他设备（甚至一些根本没有建造的机器，如巴贝奇的分析机）都将成为“第一”的候选人。[17]

威廉姆斯是对的。事实上，由于同样的原理会被许多人想到，以及将原理付诸实施的行动的多样性，这通常使得将“第一名”的殊荣完全给予某一个人或某一组人都显得缺乏说服力。如果一定要把“第一个”发明的荣誉归功于谁，那么一定是那个第一个拥有清晰原理的人或团队。他们看到了原理的潜力，努力使其获得市场的接纳，并最终令其获得充分的使用。但通常会有好几个这样的人或团队几乎同时存在。

其实即使是一个单一的发起人，由于人际互动和信息网络[18]的存在，也已经大大扩展了我描述的那个发明过程。这些交流会使发起人沉浸到相关问题和先前所进行的尝试当中，并提供原理在其他域中应用的建议以及相关的设备和技术诀窍，从而有助于将概念转变成物理实在。

科学与数学中的发明

我所描述的关于发明的逻辑结构能否被拓展到科学和数学的起源上呢？我的回答是：如果加上一些必要的变化的话，是可以的。理由是，不论科学理论还是数学理论，其目的与技术一样，都是要使其系统化。它们的结构来自于那个完成给定目的的组件系统，因此，技术的逻辑同样可以应用在它们之上。

让我用一个读者耳熟能详的科学例子来阐明这一观点。在达尔文完成贝格尔号航行后大约一年的时间里，他一直在寻找一个关于物种进化的理论，来解释诸如他在加拉帕戈斯群岛观察到的不同种类的雀类是如何形成的问题。从他的阅读和经验当中，他将一组事实和观点放在一起，这可以帮助他找到一个支撑原理：进化的时间尺度与地质时间是相符合的；个体应该是物种形成的中心元；性状的变化在某种程度上可以遗传；变异可以使物种适应缓慢变化的环境；在个体生命过程中获得的习性，可能会以某种方式促成可遗传的变化；动物饲养员可以选择他们想要的对遗传有利的特质。事实上，“我很快就认识到选择是人们成功获得动植物育种的基石。但是选择是如何被应用于生活在自然状态中的生物身上的，对我来说还是个谜”。达尔文纠结于这些不同的候选组件将如何共同建构起一个关于物种进化的解释。

在1838年，达尔文写道：“我为了消遣，碰巧阅读了马尔萨斯的人口论，并为去理解书中随处可见的生存斗争的论述做好了思想准备。这种准备来自于我对动植物生活习性的长期观察。醍醐灌顶般地，我认识到，在这种情况下，有利于生存的变异将会被保存下来，而不利的将被摧毁。结果就会形成新的物种。这时，我终于找到了一个有效的理论。”

用我的话来讲，达尔文并没有从马尔萨斯那里得来理论，他只是借用了一个次级原理：对稀缺资源的持续的竞争选择了群体中最具适应性的个体。然后他应用这个次级原理，使得他的两个主要原理之一成为可能：那些有利的适应会被选择，并被积累下来，从而产生新的物种。但对于另一个主要原理，即变异产生了一系列特征，在此基础上选择才得以进行，他还不能归纳出更好的解释，所以不得不将其作为一个前提。但是我们必须承认，将部分组合起来形成解释功能，使他找到了一个令他得以展开工作的理论。经过15个星期的艰苦思考，他完成了他的基本原理。而余下的工作，即运用所有的支持片段进行从基本原理到完整理论的细节转化，并最终使自己达到满意的过程，则用了大约20年。

科学理论化的起源说到底也和技术一样是一种链接，一种对一个可观察的给定问题与一个对此有模糊暗示的原理之间的链接。科学最后需

要用一套完整的原理再现这一切。

数学中的起源又是怎样的呢？它也是一种链接。但这次需要被证明成为某些概念形式或原理的东西（通常是一个定理）共同构成一个证明过程。设想一个定理是一轮精心建构的逻辑论证，如果它是在已被接受的逻辑规则之下建立的，那么它就是有效的。这些规则来自其他有效的数学内容，主要包括其他定理、定义以及那些构成数学有效组件的辅助定理。

一般来说，数学家“看到”或模糊地感到一个或两个主要原理，即一种概念性的想法，然后以可证明的途径提供某种整体的解决方案来进行证明，这些用于证明的方案必须来自其他公认的次级原理或定理，最后再去除每个部分有争议的部分。安德鲁·怀尔斯（Andrew Wiles）对费马大定理的证明就是使用了日本数学家谷山纪章和志村五郎的模块化和椭圆方程作为基本原理来链接他需要的两个主要结构的。

为了证明这个猜想并链接构成争论的各部分，怀尔斯采用了很多次级原理（subprinciples）。“你翻到一个页面，就会看到那里有德利涅（Deligne）的一些基本原理的简短概观”，数学家肯尼斯·里贝特（Kenneth Ribet）说，“然后你翻到另一页，多少有点巧合，你会看到赫勒高曲（Hellegouarch）的原理——所有这些都会被召来参与并在下一个念头出现之前被短暂地利用。”[19]这个整体是一种原理（概念性理念）的级联（concatenation），它们共同建构以达成目的。而每个原理构成或原理又都源于一些更早期的级联。和技术一样，每个链接都提供了一些通用的功能，一些构成争论的关键部分，它们被用于整体结构之中。

科学与数学中的原创和技术中的原创没有根本性的不同，我们不必对此感到惊讶。这种对应的存在不是因为科学、数学和技术是一样的，而在于三者都是目的性系统。广泛来说，也可视为达到目的的手段，因此需要遵循同样的逻辑。三者的构成都始于形式或原理：对于技术，是源于概念性的方法；对于科学，是源于解释性的结构；对于数学，则源于真理与基本的公理结构。因此，技术、科学和数学的产生都是通过类似的试探启发过程——基本上是通过存在于问题和能满足它的形式之间的一个链接来完成的。

发明与新的构件

我们现在有了关于“新技术是如何产生的”这个关键问题的答案了。其机制当然不是达尔文主义的。技术中的新物种并不是产生于微小变化的积累。它们产生于一个过程，一个人类的漫长过程；一个将需要和能满足需要的某个原理（某个效应的一般性应用）链接起来的过程。这个链接从需求自身出发，延伸到能够被驾驭的某个基本现象，再通过配套解决方案以及次级解决方案最终使需求得以满足，并且使其界定出了一个递归性的过程。这个过程不断进行类似的重复，直到每个次级问题消解到可以进行物理性解决的程度。最后，问题一定会被那些已经存在的片段、成分，或者那些由现存部分创造出来的片段所解决。发明是从已有之物中产生出来的。

我们现在可以理解为什么发明是如此不同了，因为每个个别的案例可以分别源于需求驱动或者现象驱动；发起者可以是一个人也可以是很多人；发明原理可能很难被想象，也可能自然地就流露出来了；把原理转化为物理组件可能很简单，也可能只有当关键的次生问题被解决之后才能进行下去。但是无论它们经历了什么，最终所有的发明共享同样的机制：所有发明都是目的与完成目的的原理之间的链接，并且所有发明都必须将原理转化成工作元件。

那么在关于新技术如何从技术集合中建构出来的问题，上述论述又能告诉我们什么呢？将本章论述和第5章结合起来，我们可以说，新构件的产生有三个可能的途径：作为标准工程问题的解决方案（阿姆斯特朗振荡器），非刻意审慎的发明（货币制度），以及新发明，即在全新的原理之下，全新的根本性解决方案（喷气式飞机发动机）。无论哪种情况，发明都是产生于那些提供了制造新元素必要功能的已有技术（现存的元素）的组合。

至此，我们依然没有完成对个体技术的解释，一项新的技术并不会就此停滞，它是会发展变化的，或者我们可以说它存在一种狭义上的“进化”过程，会呈现出一个改进的面貌，因而会建起一个家谱。如同发明自身一样，这种发展自有其独具特征性的阶段。那么技术的发展及其背后又隐藏着什么呢？

THE NATURE OF TECHNOLOGY

第7章 结构深化

技术一旦走上发展之路，各种各样的版本就会随之出现。通过“内部替换”，开发人员可以用更好的部件（子技术）更换某一形成阻碍的部件。开发人员还可以通过寻找更好的部件或材料，或者加入新组件进行结构深化。旧设计和旧原理一经锁定，就会产生新用途。

通常来讲，一项新技术的最初版本都是粗糙的。在新技术发展初期，只要它能发挥基本效用就足够了，此时，它可能只是由现有构件或者其他技术中的零部件粗略地拼凑而成。劳伦斯最初的回旋加速器就曾用了一把餐椅、一个衣帽架、玻璃窗、封蜡，还有些黄铜当配件。诞生初期的技术只能以手边可用的组件作为基础，然后再作适当调整，并适当地扩展应用范围以尽量有效地服务于不同的目的。接下来，推动者会不断把玩这个新结构，并开始制造新配件。为此，他们需要试验更好的材料、发展理论、解决问题，当然过程中会常常碰壁。总之，这是个通过逐步的、试验性的努力取得进步和发展的过程。

> 技术思想前沿
>
> 技术的两种发展机制：
>
> 内部替换（internal replacement）和结构深化（structural deepening）。内部替换是指用更好的部件（子技术）更换某一形成阻碍的部件。结构深化是指寻找更好的部件、材料，或者加入新组件。

一项技术就这样开始走上一段发展和进化的旅程。[1]实际上，这段旅程可能早就开始了。在将基本概念转化成物理形式的过程中，人们就已经在试验不同的零部件了，并一直致力于寻求改善，因此，在技术的创始和发展之间，并没有一个清晰的界限。

无论技术发展是怎样开始的，一旦上路了，技术的不同版本就会随即出现。其中一些版本出自技术的发起者，一些则来自侵入这一新领域的其他发展者，还有一些可能来自实验室或者一些寻求发展新技术的小公司。之所以这样，一方面是因为这些行动者可能会有各自界定概念的视角；另一方面则是因为，当新技术开始适应不同的目的或市场时，会发生相应的演变，这也会使技术慢慢呈现出不同的版本。[2]例如，雷达在实现了“探测飞行物”这个基本目的后，其功能就被分别拓展到探测潜艇、空中和海上导航，以及空中交通管制等许多分支领域去了。

不同的视角和应用分支为解决问题提供了不同的方案。我们可以说，是技术本身引起了解决方案的变化，是它自己表明或暗示了需要什么样的解决方案。随着时间的推移，开发者可以获得更大的空间去选择

解决方案，这时达尔文理论中的变异和选择就真的在技术中开始发挥作用了。[3]通过选择更好的方案来解决其内部设计问题，技术的不同版本将逐步得到改善。

但设计师们也通过自己的深入思考来改进技术，对此达尔文理论是无法帮助我们弄明白他们是如何做到的，也无法对技术发展之谜给出解释。在技术逐渐成熟的过程中，它们总是倾向于变得更复杂，实际上是变得非常复杂。F-35C喷气式战斗机比莱特兄弟的飞行器要复杂多。如果硬说它是莱特兄弟飞行器的变异和选择的结果，那对我们的想象力将是巨大的挑战。这里应该有什么东西是超越单纯的变异和选择而起作用的，让我们回到技术发展过程中重新审视一番，看看到底是什么在起作用吧。

内部替换

当一项技术涉及了商业或军事的议题，它的功能性就将可能受到“促逼”（pushed）：它被逼迫着给予更多功能。为了赢得竞争，投资人会寻找更好的组件、更优化的结构，不断进行组件间的调整和平衡。如果竞争非常激烈，甚至技术的边边角角都会被要求完成得尽善尽美。

但是一旦系统中的某些部件遇到了限制，那么技术（或者更准确地说，技术的基本原理）就可能无法再继续向前了。尽管开发商们希望集成电路可以排布得更密集、容纳更多器件，但是在该项技术发展的过程中，总会遇到限制出现的那一刻，比如光刻法最终会受限于光的波长。在雷达发展的早期，人们不断尝试以更高的频率发射信号，以便能够更准确地识别目标。但是在如此高频的情况下，信号发射源将无法维持稳定的功率。一项技术在遭遇局限性后只能就此停步。

限制性障碍令技术投资者大为恼火。[4]为了进一步发展，遭遇到的每个瓶颈都必须得到认真对待。[5]令人沮丧的是，限制是不可避免的。假设一个设计没能接近操作的极限，它的效率就不是最充分的，它会继续被要求表现得更充分。

通常，开发人员可以通过更换形成阻碍的零部件（一个次生技术）来克服局限。这种置换可以通过以下方式实现：采取更好的设计或更深思熟虑的解决方案，或者天才地盗用竞争对手的思路等。另一种方法是，用不同的材料，比如强度更大或熔点更高的材料进行替代。喷气式发动机在开发的数十年间，就一直在不断改用更强、更耐热的合金零部

件。事实上，开发者经常寻找的并不是更好的部件，而是这个部件恰好能提供的一个现象。因而，开发者在寻找化学性质相似的材料时最在意的是，哪一种材料在被利用时会出现更有效的现象。的确可以这样说，大多数的材料科学是通过理解材料性能来寻求现象的改善的。

当然，一部分组件的改进需要其他组件也进行适应性地调整。这个过程需要对整个技术再次进行平衡，甚至可能需要重新考量技术的整体结构。20世纪二三十年代期间，木结构飞机被替换成了金属框架，那使得整个飞机设计都必须重新加以考量。

关于通过内部替换进行改进的过程看起来好像已经很完整了，但是按照我们的递归原理，这个过程也应该是递归性的。技术的改进过程应该伴随着构成次级组件以及次次级组件的零部件的置换过程，也就是说，我们需要将作为客观对象（object，其实更像一种有机体）的技术的发展视为一个在所有层级上的所有组件都同时发生改进的过程。

除此之外，一项技术的发展不仅需要那些内部的、直接的努力，还需要目标技术以外的“外部”改善。这是因为许多技术组件都来自于其他技术。几十年以来，航空仪表和控制原理都不断得益于其“外部”的电子领域的发展。技术的发展依托于其构成组件所需的外部发展。

结构深化

内部替换部分地解释了技术为什么会随着发展而变得越加复杂。我们可以将之理解为，替换构件平均来讲要比它们所取代的部件复杂。但那不是重点，真正起作用的另有原因。我将其称为结构深化。[6]

开发人员可以围绕着技术障碍去寻找更好的部件或更好的材料。他们也可以通过加入新的零部件或是添加进一步的零件系统去消除障碍。这时，解决障碍的方法不是调换旧零件，旧零件实际上已经被保留下来了。在极限圈定的范围内，会有其他零件或集成件被添加进来，以辅助已有的旧零件完成工作。因此，当喷气式发动机被要求在更高的温度下运行而导致涡轮叶片开始在极限温度软化时，开发者采取了在叶片旁边添加通风系统冷却叶片，或者在叶片内部添加循环冷却系统的方法。当早期雷达系统接收的回声信号与发射信号因同步而被淹没时，开发人员加入了一组零件（双工器），这可以使发射器有零点几秒的短时关闭，这样回声就可以被清晰地接收到了。

为了突破局限而不断加入次级系统，技术因此发展得越来越精致。技术结构就是这样不断被“加深”或者不断地被设计得更为复杂的。

技术从而变成了重重叠叠的复合体。

驱动技术复杂化的动因不仅在于为完成目标功能而被迫去克服技术限制的尝试。一项技术不仅需要自身运行良好，还需要在外部环境发生变化时也可以应付自如。也就是说，它应该可以安全可靠地应对一系列任务。而在这个过程中，极限无处不在。所以我们可以说，为了克服各种极限，一个技术还需要主动增加次级系统或次级模块以完成如下目的：（1）加强基本性能。（2）对改变或异常进行监视并作出反应。（3）去适应更广泛的任务范围。（4）增强安全性和可靠性。

这一观点不仅限于技术的统筹层级。由于它的次级系统或集成模块本身也是技术，也需要发展，需要被促逼着加强技术总体的性能，所以，通过递归性过程，主动改进过程也会发生。设计者会寻求打破极限的方法，根据上面（1）～（4）的原理加入次级系统从而加强性能、对环境的改变作出反应、适应更广泛的任务范围、增强可靠性。新加入的集成块或子系统反过来也会被促逼着趋近它们自己的操作极限。设计者将进一步加入次次级系统（sub-subsystem）来打破这些极限。这个过程持续进行着，集成模块围绕提高主模块的工作性能工作，其他次级模块又围绕着集成模块工作，还有其他模块再围绕着这些次级模块工作。性能在系统的所有层级上被提高，技术结构的所有级别都将变得更为复杂。

我们可以从燃气涡轮飞机发动机这个例子来看这个递归性的结构深化过程。弗兰克·惠特尔的发动机原型是用独体压缩机来供应压缩空气进行燃料燃烧的。这是一个径向流动（radial-flow）压缩机：它吸入空气，然后通过快速“旋转”进行压缩。惠特尔熟悉这种压缩方式，之所以选择这种技术，是因为它是达到目的的最简单的设计。但是随着对更好性能的需求，设计师被要求采用更好的组件，即轴流式压缩机来替换径向流动压缩机。它是一个巨大的风扇，其空气流动方向平行于传动轴。但在单轴压缩机阶段，增加压力供应比的极限也只能达到大约1.2：1。为了达到更优越的性能，设计师同时使用几个压缩机，并最终将它们按顺序排列组装在一起。但是这个压缩系统的操作需要既能适应高纬度的稀薄空气，又能适应低纬度的高密度空气，并且还要适应不同风速的操作环境，因而设计师增添了导叶（guide-vane）系统来控制吸入的空气。系统因而被精致化了。反过来，导叶系统现在又需要一个可控集成

块来感知环境条件，从而对叶片进行相应的调整。这是进一步的精致化。但是现在输出的高压空气会以意想不到的冲击波回冲到压缩机中，这又成了另一个主要的障碍。所以压缩机又需要安装防喘振放压阀这一次级系统对此加以控制，并且要对该系统进行更进一步的精致化。防喘振系统又需要更灵敏的传感、控制系统。此外，还要有更多的精致化过程。

以这样的方式，一项技术（例如，压缩机）在性能和使用范围上得到了极大改善，但这是要付出代价的。随着时间的推移，它只能背负着越来越多的老旧的子系统和次级子系统才能正常运转，它要不断处理各种异常情况、被迫扩大应用范围、为防止或应付失败而进一步提供大量冗余设计。

在喷气式发动机的案例中，当技术被促逼时，也会发生为提高性能而主动进行改进的情况。例如，为了提供空战条件下要求的额外的推动力，特意添加了补燃室这个集成件；为了防止发动机起火，特意加入了复杂的烟火警探测系统；为了防止通风口结冰，特意加入了除冰组件。此外，专门的燃油系统、润滑系统、可变尾喷系统、启动系统也都是因此被加入进来的。所有这些反过来又需要控制、传感、仪表测量系统及其子系统。如此一来，飞机性能确实被提高了，现代飞机的发动机动力比惠特尔最初的喷气式发动机要至少高30～50倍，但它们也更复杂了，惠特尔1936年的发动机包括一个移动的涡轮增压机以及几百个零件，而它的现代版已包括22 000个零部件了。

通过结构深化来改进技术的过程是缓慢的，飞机涡轮增压汽油发动机的改进前后共用了几十年。原因在于，改进过程中不仅要识别新的集成模块及其问题，还必须对它们加以实验、论证，以及对即将涵盖它们的新系统进行重新平衡和优化，而这些都需要时间。

经济因素也在这个过程中帮助控制改进的时机。如果竞争激烈，改进就会加速，如果缺乏竞争，改进就会慢下来。开发人员即使意识到明显的技术进步，也不一定会予以采用。不论竞争压力是否出现，任何时候发生的技术改进都会被认真进行选择，某个技术上可行的新改进必须是在经济上进行考量，认为值得进行整体的重新设计后，才可能被采用。

就这样，随着新的改进被选择性地采用，技术一点一点向前蹒跚发展。如果遇到某个限制性的阻碍，发展会缓慢下来，致使整个过程显得时断时续。总之，技术的发展深深依赖于结构的深化。

结构深化对技术进步的作用往往是巨大的。但是，随着时间的推移，通过不断加入系统、子系统而获得更高的性能之后，新技术也会遭遇硬壳化（encrust）。硬壳化对于一个方法和设备可能不算什么，一旦开发成本、费用已完成分摊，导致的结果可能仅仅是材料或空间使用成本的提高或权重的增加。但对于所谓“非技术性的”目的系统，负担则可能相当沉重，如军事组织、法律制度、高校管理以及文字处理系统都需要通过不断加入子系统或子零件来赢得性能上的改进。只需设想一下税法的复杂性逐步增加的过程就可见一斑，而它只不过是法律制度体系中的一个小分支。这些以复杂性和官僚主义的形式表现出来的“改进”成本是无法分摊的，即使环境已经不再需要它们了，它们还将继续存在并且很难消除。

锁定与适应性延伸

我所描述的内部替换和结构深化这两种发展机制，将作用于技术的整个生命周期。起初，一项新技术被小心翼翼地、试验性地开发出来。后来，当它能够服务于某一特定目的时，它就迅速发展起来，并成为标准工程的一部分。当然，同时，达尔文机制开始起作用，它从许多内部的改进中去选择（通常是通过设计师对前辈的借鉴）更好的解决方案。

技术思想前沿

在新旧原理更替的过程中，旧原理往往已经被锁定了，有4个原因导致旧原理通常会存在较长的时间：

- 经过精致、繁复的过程之后，已经成熟的旧原理会表现得比它的新对手好。
- 采用新原理可能意味着改变周围的结构和组织，因为成本过高，所以可能不会被实现。
- 从业者不认可这个新原理带来的愿景或承诺。
- 新原理将使旧知识过时，它在潜在的新原理与安全的旧原理之间制造了一种认知失调以及情感上的不匹配。

当调换部件和结构深化都不能再为提高性能做什么的时候，技术就成熟了。这时候，如果想取得进一步的发展，则需要一个全新的原理。

但新原理不能说出现就出现，即使出现了，它想要取代旧原理也不是那么容易的。旧设计、旧原理往往已经被锁定了。

这是为什么呢？一个原因在于，经历如此精致、繁复的过程之后，已经成熟的旧技术会表现得比它的新对手好。新对手们可能潜力无限，但此时它们还处在婴儿期，因而无法马上追赶上那些老道的旧技术。它们的成熟不可能一蹴而就，因此旧技术存在的时间通常比预料的长。

还有一个理由使得旧原理存在的时间要比应然的长，在经济领域即是如此。即使新的原理发展得很好，表现也比旧原理好，但是采用它可能意味着改变周围的结构和组织。由于这样做成本过高，因而可能不会被实现。在1955年，经济学家马文·弗兰克尔（Marvin Frankel）很想知道为什么兰开夏郡（Lancashire）的棉纺厂没有像他们的美国同行那样采用更先进、更高效的机器。[7]他发现，如果在英国设置新机器，的确会更有效率。但是这些新机器很重，如果安装它们，那么安置旧机器的维多利亚时代的砖结构就会被拆除。就这样，“外部”的组件或经营锁定了内部机械，导致兰开夏郡的棉纺厂没有改变。

还有一个原因是心理上的。旧原理可以持续下去是因为从业者不认可这个新原理带来的愿景或承诺。首创不仅仅是一种新的做事方式，而且也是一种新的看待事物的方法。

还有新的威胁：新原理将使旧知识过时。事实上，有些新原理已经被推广过了或业已存在，但是它们被从业者所摒弃。有时候并不是因为人们缺乏想象力，而是因为它在潜在的新原理与安全的旧原理之间制造了一种认知失调以及情感上的不匹配。社会学家黛安娜·沃恩（Diane Vaughan）谈到过这种心理失调：

> （当我们作为人类去处理情况时，我们会使用）一个由一套假设、期望和经验构成的综合的参照系。所有事物在这个框架的基础上被认知的。这个框架会变成自证实的，因为，只要我们能够，我们总是倾向于将这个框架置于经验、事件之上，创造符合它的变故和关系。我们对那些不符合这种参考框架的事物往往采取忽视、误解或否认的态度。结果是，我们通常能找到我们所要寻找的东西。这个参照系是不易被改变或铲除的，因为，我们看待世界的倾向，与我们怎样看待和界定自己与世界的关系是密切相连的。因此，在我们与这个参考框架保持一致的过程中，是有自己的既得利益的，否则我们自己的身

份识别就将遭遇危险。[8]

它和新原理的关系也是如此。新的和已被接受的解决方式之间的距离越大，对传统方式的锁定就越牢固。因此，迟滞现象（hysteresis）是存在的，即对变化的一种延迟反应。新技术被非常成功的旧技术所阻碍，技术上的转换既不容易也不顺畅。

这种对旧有的成功原理的锁定所引起的现象，我称之为自适应延伸（adaptive stretch）。[9]当一个新的情况出现或要求在其他领域应用时，人们更容易想到用旧技术或旧有的基本原理加以解决，并且会通过“拉伸”它来涵盖新的环境。

技术思想前沿

自适应延伸：对旧有的成功原理的锁定所引起的现象。

在实践中，这意味着将旧技术的标准组件重新配置以应用于新用途，或增加更多的集成模块来完成新的目标。20世纪30年代，喷气式发动机的高速高海拔飞行本来可以提前好几年实现，但是设计者当时还不熟悉燃气涡轮机的原理。于是，当军用飞机被迫在空气稀薄的更高的高度飞行时，他们适应性地延伸了当时的技术：飞机活塞发动机。这迫使活塞发动机要打破极限[10]：这些极限不仅包括高海拔氧气稀少，还包括能够将氧气以足够快的速度泵入汽缸的极限，它受氧气在四冲程发动机系统中燃烧和处理的速率所限。增压器和其他系统的深化被加入进来，用以进行高压快速泵入空气。活塞螺旋桨推进器的活塞部分被独具匠心地精致化了，它被延伸了。较难延伸的是螺旋桨推进器。如果长时间在空气稀薄的高空飞行，螺旋桨推进器就会无法咬合。如果它被促逼着进行更高层次的改进，它就可能走向超音速。如果它被扩大了，拥有了更大的直径，螺旋桨尖部的运动会更快，同样也会走向超音速。一个根本性的极限已经达到了。

这是一个很典型的案例。在发展的某种程度上，旧原理变得很难再进行扩展延伸。这时新原理就有了向前发展的立足点。当然旧原理还会在附近徘徊，但是它已经变成为某个特定目标服务了，而新原理已开始了精致化过程。

我在本章和第6章所描述的新原理的起源、结构深化、锁定以及适应性延伸等现象组成了一个自然周期。一个新的原理出现、开始发展、

陷入局限、其结构不断被精致化。环境结构和专业熟悉度被锁定在原理及基础技术当中。新的目的和变化的环境出现时，它们通过延伸锁定技术进行适应，从而产生了进一步的精致化。最后，已被高度精致化的旧原理已经超出了它能承受的极限，因此将让位于新的原理。新的基本原理可能更简单，但在适当的时候，它会自己变得精致化。

这个周期循环往复，有时简单性会切入到精致化的过程中。精致化和简单性就这样来回交替，直到精致化进程随时间的推移而到达它的边缘。

这可能会使我的读者受到震撼，因为我正在描述的整个循环与托马斯·库恩就科学理论发展所提出的周期描述非常相似[11]。库恩所说的周期肇始于一个新原理（他称之为范式）取代旧原理的新理论模型。新范式之后被应用于许多样本，从而被接受，并进一步在库恩所谓的常规科学中被精致化。随着时间的推移，不适合那个基本范式的例子（异常）逐渐树立。范式会通过伸展来适应这些，但是当异常进一步建立起来，范式就会越来越勉强。当且仅当一个更新的、更令人满意的解释，即一个新范式出现后，旧范式就崩溃了。

如果想要展现这种对应性，我可以用库恩的术语来重新阐述技术周期。但是用我在本章使用的术语来重述库恩理论将更有趣。我们可以说一个理论通过面临新的事实和被迫进行新的应用而被促逼着。它的组件可能需要被替换，可能需要更准确的定义，很多已有的构件可能需要重构。当遇到限制（用库恩的话说就是异常）的时候，许多方面就会发生特殊情况：这时真正在工作区附近徘徊的系统就会进行精致化，从而应对这种可感知的限制。这个理论建立的基础是通过加入次级理论来处理困难和特殊的情况。例如达尔文的理论就必须加入次级论述来解释为什么有些物种会展现利他性。

理论的发展就是进行精致化的过程，它加入附录、完善定义、添加补编以及特殊结构，其目的在于将所有特例纳入考虑范围。如果特例不太符合理论，理论就会延伸，它就会加入相当于“本轮”[10]的东西。最终当面临足够的变异时，它的“功效性”就降低了，此时就要寻找一个新的原理或范式。当现存范式无法再被延伸时，新的结构就开始形成。库恩的循环开始往复。

我重申这种一致性并不意味着科学和技术是一样的。它只意味着，由于科学的理论体系是目的性的，因此它们遵循同样的逻辑。科学也会发展，碰到局限性，进行精致化，并在适当的时候寻求替换。无论是科

学还是技术，其发展变化的逻辑是相似的。

本章我们一直在探索技术发展的过程。但我们要注意，它适用于技术所有的部分。新的技术被促逼着陷入局限，并通过优化零部件和深化结构来获得提高。这个过程也同样适用于所有的零部件。发展是非常典型的内部过程。整个技术和它的所有部件并行着。

THE NATURE OF TECHNOLOGY

第8章 颠覆性改变与重新域定

域并不是若干单个技术的简单相加，它们是连贯的整体，对经济的影响也更大。任何新域，都产生于一个已存在的域——母域，而且参与者开始很少能意识到会发生“颠覆性改变”。但随着理解的深化和实践的固化，新域会慢慢忘记它的母域而横空出世，甚至会极大地提升国家竞争力。

前两章回答了关于个体技术是如何形成和发展的问题，那么，技术体又是如何形成和发展的呢？我将在本章研究这个问题。

许多需要解释的问题似乎都讲过了。如果我们只将技术体（或者如我所定义的“域”）看作是构成它们的技术和实践，以及那些随时间不断产出的新技术的总和，那么，好像也不必再用单独的一章来讨论它了。

域实际上并不是单个技术的加和，它们是连贯的整体（coherent wholes），是关于设备、方法、实践的族群，它们的形成和发展具有与个体技术不同的特征。它们不是被发明出来的[1]，而是通过类似结晶的过程，从一套现象或者一种新技术的可能性当中浮现，并由此开始建构起来的。技术体发展的时间跨度不是几年，而是十几年、几十年。比如，数字域出现于20世纪40年代，但是直到现在，它依然在扩展延伸。技术体的发展不是某个人或一小部分人能够推动的，而是需要众多的相关利益群体的共同参与才能实现。

域对经济的影响也比单个技术对经济的影响要更深刻。比如，1829年的火车，即使它在19世纪30年代进行了改进，在经济上也没有显现太多的回应，但是当包括铁路在内的一个技术的域开始形成的时候，经济不但开始回应，而且随之发生了巨变。实际上，我想说的是，经济并不是采用（adopt）了一个新的技术体，而是遭遇（encounters）了一个新的技术体。经济对新的技术体的出现会作出反应，它会改变活动方式、产业构成以及制度安排，也就是说，经济会因新的技术体而改变自身的结构。如果改变的结果足够重要，我们就会宣称发生了一场颠覆性改变。

域是如何进化的

让我们深入到本章的主题中去吧。域究竟是如何形成并发展的呢？

域的发展，如我所说，主要是围绕核心技术联合而成的。随着计算机的诞生，其支撑技术，如读卡机、打印机、外部存储系统、程序语言等，就开始在它周围聚集起来。另一方面，则由各个学科领域（产生于现象族群以及相应的理解和实践）的形成和发展所构成。比如，电学和无线电工程的建立就发端于对电子和电磁波的理解。

技术思想前沿

域的形成有两种模式：一是围绕核心技术联合而成的，一是从一个现象簇中建构起来的。

无论域是从新技术中“结晶”出来的，还是从一个现象簇中建构起来的，它们都产生于一个业已存在的领域。这是因为它们的起源部分和最初理解必然是有来处的，就是它们所谓的母域。例如，辅助计算并不是从处理器或数据总线中产生的，而是来自于20世纪40年代的真空电子管的元器件及其实践领域。

在开始时，几乎很难将一个新领域称为一个域。它只是将一堆粗浅的理解和方法松散地堆积在一起，此时它能提供的东西很少。当时的经济活动在对待这个新生事物时，一般也都采取保守的态度。但当经济需要时，则会浅尝一下，或者从中作出一些选择，可能就此会形成一个杂交组合的域，也就是说，既有新域的成分，也有母域的成分。早期的英国铁路，斯托克顿-达灵顿专线其实就是在铁轨上跑马车。新域的部件经常都是作为旧域的补充而被征用的。早期的蒸汽机就曾在枯水期将水引向高处蓄水池以帮助驱动水车。

初具雏形的域在很大程度上属于母域的一部分。基因工程[11]最初只是其母域（分子生物学和生物化学领域）的一个很小的分支。它主要源于20世纪中期的数十年间，对细胞及其DNA制造蛋白质的机理的那些理解。到20世纪70年代早期，生物学家开始逐渐理解了特定的酶（限制酵素）如何在特定的地方分开或者切断DNA的，而其他的酶（连接酶）又怎样将DNA片段连结在一起，以及另外一些酶怎样控制DNA的复制过程。这就意味着他们已经开始理解细胞是如何利用“天然技术”复制基因，并指导它们生产特定蛋白质的。慢慢地，他们开始捕捉这些现象，并尝试在实验室里进行人工复制。生物学家在那时已经开发出一套技术，来完成一些以前只有大自然才能够完成的任务。

就是这样出发，一个新的域开始慢慢浮现。

任何域在其产生的早期，都有类似这样的经验性感觉。但是参与者在这个阶段通常很少能意识到会有一种新的技术体出现。参与者会认为自身正在解决母域中存在的某些特定问题，但最后，新的集群终将获得自己的词汇和思维方式。随着关于它的理解的深化、它的实践的固化，它的每个组件都会经历如我在第7章中谈到的那类发展过程。接着，新

域会慢慢忘记它的母域，并伺机表明自己已经独立了。这时我们就会意识到，一个新域凭借自身的力量诞生了。

技术思想前沿

域的生命周期：

- 诞生。解决母域中的特定问题，在理解和实践中固化、发展。
- 青春期。解决发展中的阻碍，产生可行的技术并应用于市场。
- 成熟期。市场由狂热走向冷静，新的域以自己的方式深入地影响经济，进入稳定成长阶段。
- 晚年。鲜有重要理念产生，有些域会被取代，但大多数还得以存在并服务于人类。

一个新域在刚开始形成的时候，可能只是一个即将装满技术和实践的工具箱，但在这个工具箱里，最关键的构件在功能上可能还很弱，或者甚至根本还没出现。这样域的发展会遇到一种意想不到的阻碍，这个阻碍（历史学家托马斯·休斯将其称为“反向凸角”，意指在发展和前进唾手可得之时遇到的阻碍）会引起参与者们的注意，于是所有的努力便会向那里集中。如果努力足够，再加上一点运气的话，域的发展可能就会适时地实现突破。

如果域突破的这个阻碍非常重要，那么就可能产生一项可行的技术，即一个可以实现重大商业应用的手段，或一些可供继续建构未来技术的关键性元素。那些将基因工程带入了生活，从生物的一个分支转化为一个新的生物技术领域的突破是斯坦利·科恩（Stanley Cohen）和赫伯特·博伊尔（Herbert Boyer）在1973年对DNA进行重组实验时实现的。当时科恩博士一直在研究细菌中的某种质体，那是一个微小的DNA循环单位。博伊尔则致力于研究限制性酶，那是一种可以切断DNA链条的“分子剪刀”。他们将两种方法合在一起，开发出了一种技术：从一个生物体（在他们早期实验中用的是一种属于非洲爪蟾的青蛙）上切下一段基因，粘到一个质粒当中，然后将质粒插入到一个大肠杆菌中，它就会带着它的外来基因迅速地进行自我复制。这时细菌就会作为一个小型工厂，生产那个外来基因的蛋白质。

DNA重组技术最初是粗糙的，还因其可能具有非自然性地创造生命的潜在可能性而引发巨大的争议。抛开争议不谈，这项技术突破了两个

基因技术发展的极限：从生物体剥离基因和人工制造蛋白质。

正是由于这样的“扩展人的能力的技术”，新兴的域慢慢进入它的青春期，甚至跨入成年阶段。它现在可以发挥作用了。技术开拓者们开始聚集，市场上会逐渐出现一些小公司。接着，大量的改进随之跟进，机会也随之出现。一个行业开始成长，新的领域开始活跃起来，记者们推波助澜，投资蜂拥而来，争抢非凡的利润前景。基因泰克（Genentech）是第一个基因技术公司，当它宣布上市时，我们曾见证了它的股票价格在首个交易日的20分钟之内从其发行价的35美元飙升到了89美元的过程。在这种情况下，上市的许多公司除了某个理念，其实没有更多的东西可以售卖。基因泰克于1980年上市，但是直到两年后，它才真正拥有并开始销售它的产品——人工胰岛素。

经济学家卡洛塔·佩雷斯（Carlota Perez）曾经探讨了技术革命爆发的那个特定阶段。[3]他指出这种情况通常会引起投资的热潮或崩溃。在基因技术发展的早期阶段，并没有发生投资崩溃的情况，但是崩溃的例子在历史上绝对不在少数。

19世纪40年代中期的英国，到处充斥着对铁路的狂热。“这个国家简直就是一所铁路疯子收容院。”洛德·考克伯恩（Lord Cockburn），一位苏格兰法官，曾这样描述过当时的场景。这种疯狂带来了不可避免的金融崩溃。到1845年时，铁路狂热达到顶峰，街边小贩们都在售卖临时股（股票被分成小份），几乎天天都有新的开发计划在冲击着一个又一个不知名的小镇。然后突然间，泡沫破裂了。经济恐怖周开始了，1847年10月16日，铁路股票市值锐减85%，许多银行被迫关闭，英国经济霎时走到崩溃的边缘。当然并不是所有的新兴领域都会伴有金融危机。通常，发生危机的情况都是由于存在一种空间争夺。机会只有那么多，要修建一条铁路，要么在曼联和利物浦之间，要么在伦敦和伯明翰之间，两者只能选一。当然即使存在危机，新的域还是有机会存活下来的。

随着进一步的成熟，新域会开始以它的方式深入地影响经济。最后，它会进入一个稳定成长阶段。早期的竞争狂热结束了，大多数小公司已经消失。幸存者成长为大企业。新阶段有了不同的气氛，它冷静、努力、充满信心并成长稳健。新技术已经找到了它正确的位置，并成为经济发展的潜在部分。这个时期可以持续几十年，是技术成长的稳定时期。英国的铁轨长度从泡沫高峰时期的3 457公里发展到65年后的33 796公里，在此期间它的发展是英国经济增长的动力。

随着时光的流逝，域到达了舒适的晚年。仍然会有新专利被承认，

但是这时鲜有重要的理念产生，昔日迷人的域不再灵感迸发。某些域在这个老年时期会被新域取代，然后慢慢枯萎——运河在铁路到来之后就静静地离开了。但是大多数域还得以存在。如果新技术需要，它们会变成老仆，忠实、好用，并且这在很大程度上被视为是理所当然的。和100年前一样，我们仍然在使用桥梁、道路、下水道系统和电气照明。旧时代的技术还在坚持，[4]甚至能使锈蚀的尊严发出光芒，但是我们在使用它们时，却通常视而不见。

并不是所有的域都能完成整个循环，条理清晰地度过青春期、成熟期和晚年几个阶段。一些域每隔几年就会通过重构自身（改变它们的特性）打破周期，亦即，它们可能会变异（morph）。[5]

当一个域的关键技术发生了根本性的改变时，它就会发生变异。当晶体管取代了电子管，电子技术就变异了。更常见的变异发生在它主要的应用领域发生变化的时候。在20世纪40年代，计算技术在很大程度上是以辅助战时科学计算为目的的。到了20世纪60年代，大型主机时代，计算技术就以商业和财会计算为目的了。到了80年代，计算技术又改成要为办公室的个人电脑服务了。到了90年代，辅助计算则主要参与互联网服务和商业。现今它又以所谓的网络智能技术为基础了。价格低廉的微型传感器构成的网络现在可以提供看、听、相互进行无线信息传送等功能。当它们被安装到卡车、货架上的产品、工厂的机器上时，它们就能“交流”并集体采取智能行动了。计算技术，如我所说，又一次变异了。

这样的变异常常发生在旧域发展达到顶峰的时候，它会迫使那些域的追随者们改变他们关于这个域的总体看法。但即使是这样的彻底改变，域的基本原理还是会保持一致；计算技术依然保持将所有对象的操作都还原成数字的方式来完成。这就好像同样的演员在后台换了服装再到舞台上出演不同的角色一样。

除了变体，域还会抛出新的次级域，这些新的后代通常有多种来源。互联网或者我们称之为信息技术的这种更一般的域，都是计算和通信域的子孙。它们是高速数据操作技术和高速数据传输技术联姻的结果。这些母域还在，但它们的子孙已经独立了，并且已经向分支领域迈进了。这种变异倾向以及产生新的次级域的倾向使技术体带有了一丝生命的性质，它们不是固定的组合，而是变成了一种微型生态：零部件和实践在任何时候都必须匹配得很好，并在新元素进入时不断作出相呼应的变化，还要不时抛弃一些有不同个性的小的次级群落。

这一切都意味着，域，以及任何暗示技术体的东西，从来不能被条理清晰地定义。它们补充或失去元素，从其他领域借用或交换零件，不时抛弃新的次级域，甚至当它们都相对较好地被定义和固定下来时，它们的角色和实践也会随时间和地点的不同而变化。构成计算技术的元素在硅谷和在日本一定是不同的。

经济的重新域定

当这一切发生的时候，当技术体出现并发展时，会对经济有什么影响呢？

当个体技术出现并发展时，其在经济学上是清晰的。一项新技术，例如贝塞麦炼钢法，一旦被采用并在钢铁生产商中传播，就改变了商品和服务的经济模式。和坩埚钢相比，贝塞麦炼钢法生产的钢便宜了很多，所以钢在经济中被更多地应用了。使用钢的产业（铁路、建筑、重型机械）都会直接从中获益，从而导致其成本和所能提供给客户的东西也改变了，随后应用这些产业的产业也会受到影响。这样的展开具有典型性。就如同牵拉一根丝会引起整个蜘蛛网的伸展和重塑一样，一项新技术的到来会引起经济中的价格和生产网络在各行各业伸展、重塑。[6]

类似的过程也适用于技术体。19世纪50年代，铁路的出现不仅使乘坐美国的交通工具的价格更便宜了，也使地区经济中依赖运输的部分获得了调整。[7]在美国中西部地区，有一些产品，例如进口物品和必须从东海岸运来的高运输成本的制成品，这些产品的价格因为采用铁路运输的关系而变得便宜起来。中西部的其他产品，如小麦和生猪，则涨价了。它们的供应商转而从东部购买了。为铁路工业提供支撑的工业也改变了。美国的钢铁生产在1850—1860的10年间，从38 000吨飙升到180 000吨，产量的巨大提升鼓励了大批量的生产方式被采用。反过来，小麦、制成品和铁在价格和可用性方面的变化也影响着依赖这些产品进行生产和销售的无数工业领域。所以，我们可以说，如同个体技术一样，技术体也引起了经济模式的扩展性调整。

虽然这样考虑问题非常有效，但它并不完整。要考察技术体，仅仅作些调整是不够的。想知道它是怎么回事，还得回到标准观点。经济学假设新技术是被“采用”的，它们被接纳并应用于经济中。对于个体技术当然是这样。钢铁生产商采用贝塞麦生产过程，他们的生产能力也相应地起了变化。但这并不能很好地描述复数的技术的情况，如计算或铁路

技术。我更倾向于认为工业、公司、商业运作等经济要素并不是“采用”了一个新的技术体，而是“遭遇”了它。正是由于这种遭遇，才产生了新工艺、新技术、新兴产业。

那么，这是如何发生的呢？可以将新的技术体看作是它的方法、设备、理解和实践的总体，然后将一个具体产业看作是它的各组织机构和商业过程的组合，以及它的生产方法和物理设备的组合。它们都属于我先前定义的广义的技术。这两种技术一个来自新的域，另一个来自某个特定的行业，它们聚集起来并相互遭遇，其结果是产生了新的组合。

因此，当银行业在20世纪60年代遭遇到辅助计算技术时，我们可以说，记账和会计活动采用了辅助计算技术。但是再具体些，银行业务并不仅仅是采用辅助计算。为了更加精确，记账的某些活动（记账程序和过程）与辅助计算活动（特殊的数据输入程序和一定数值的、文本的程序算法）合并在一起，共同形成了新的功能——数字化会计。结果是来自两个域的一个混合程序成为银行业新的计算程序。顺便说一句，这样的混合也可以称得上是采用。如果近距离观察，采用就是一个待采用领域和来自于新的可能域的功能之间的混合过程。

事实上，如果以这种方式形成的新的组合足够强大，它们就可以创建一个新的产业，或者至少一个次级产业。在银行业务电脑化之前的几十年间，它可以设计简单的期权和期货，即一种允许客户可以在未来某个时间点以一个固定的价格来购买或出售某些东西的合同。例如，这种合同允许农民在爱荷华州种植大豆，然后在6个月后以固定的价格8.40美元/蒲式耳出售，不管那时的市场价格如何。如果价格高出8.40美元，农民可以在市场上销售；如果价格低于指定价格，则农民可以行使期权，用购买期权合约的方式锁定利润。合同的价值是从实际的市场价格导出的，因此它被称为一种衍生工具。

在20世纪60年代，如何能赋予衍生合同一个合理的价格还是一个尚未解决的问题。这在经纪人当中有点黑色艺术的味道，它意味着不论是投资者还是银行，在实践中都不能对此充满信心。但在1973年，经济学家费希尔·布莱克（Fischer Black）和迈伦·斯科尔斯（Myron Scholes）解决了期权定价这个数学问题，而这确立了一个行业可以依赖的标准。在那之后不久，芝加哥贸易委员会创办了期货交易所，衍生产品市场就此启动了。

我们不能肯定地说衍生品交易采用了计算技术，那会贬低这个过程及其后果。毋宁说，是计算技术使衍生品交易得以精确地实现了。对各

个阶段（从迅速收集和存储财务数据到计算衍生产品价格的算法，再到解释交易和合同）来说，计算技术都是必要的。所以更准确地说，是衍生品交易的元素遭遇了计算技术，产生了新程序，进而建构了数字化衍生品交易技术。于是，衍生品交易活动大规模发生了。

事实上，实际发生的事有过之而无不及。银行和计算技术相互遭遇后共同创造的不仅是新的活动和产品，更重要的是一套对金融可能性进行的新的“编程”——一个具有操作可能性的新域。在适当的时候，金融工程师会开始为特定目的将期权、交易（交换现金流的合同）、期货和其他基本衍生物组合在一起，来抵御受未来的大宗商品价格的变化或者外汇波动带来的风险。一套新的经营活动就此出现了。这次金融和计算之间的遭遇实际上创造了一个新行业：财务风险管理。

这是一个创造性的转变过程。在很长一段时间里，这类金融风险都是管理的一个问题，最终它在计算技术和数学领域得到了解决。结果是它成为金融业中还在继续发展的、具有创造性的一项新技术，以及银行金融业务在计算机领域中的一次重新域定。

技术思想前沿

经济的重新域定，是指已有产业去适应新的技术体，从中提取、选择它们所需要的内容，并将其中部分零部件和新领域中的部分零部件组合起来，有时还会创造次生产业。

一般来讲，经济的重新域定就是令已有产业去适应一个新的技术体。但是它们这样做的时候不能只是采用新技术体，还需要从新的技术体中提取、选择它们所需要的内容，并将其中部分零部件和新领域中的部分零部件组合起来，有时还会创造次生产业。而当这一切发生的时候，域反过来也要作出呼应，它会增加新的功能，以更好地适用于那些会用到它的产业。

整个过程在经济中是不均衡的。当不同工业、商业和组织遭遇到新技术，以及以不同的方式和不同的比率重新配置时，会表现出非均衡性。它因小规模经济活动中的变化而向外传播，使商业组织和制度发生变化，最终使社会自身发生变化。一个新版本的经济缓慢地生成了。域和经济互助式地进行共适应（co-adapt）和共创新。

我们将这样的共变（mutual change）和共创（mutual creation）

过程称为“颠覆性改变”。经济领域中的每个时代都是某种模式，是在商业、工业以及社会中能够达到逻辑自洽的一套结构，这套结构是由在当时占有主导地位的域来确定的。当新的技术体，如铁路、电气化、大规模生产、信息技术等渗透到经济当中时，旧结构可能崩溃，新结构便取代了它的位置。一度被认为理所当然的产业被废弃了，新的产业诞生了。旧的工作方式、古老的惯例、旧的行业开始显得古怪，工作和社会中的制度安排开始重构。许多事情在经济中保持不变，但是也有许多事情将永远地不同了。

> 技术思想前沿
>
> 域和经济的共变和共创过程，称为“颠覆性改变”。

颠覆性改变只发生在对经济有改变的重要的域之中的说法是不对的。其实颠覆性改变也发生在不那么重要的域当中，例如，塑料注塑成型技术进入经济领域便导致了小规模的变化。因此，在任何时候，很多颠覆性改变会重叠、互动，共同改变经济。当新的技术体一同进入经济领域并发生作用时，它们就形成了相互一致的结构，并在经济中有了一个大约一致的模式。每个模式的出现可能都会显得很突然，然后会被锁定一段时间，接着再适时地成为下一次颠覆性改变的基础。这样的“奠定”过程，就如同一个地质层覆盖在以前所有的沉积层上面一样。

经济中的时间

所有这些新技术的展开、经济的调试，都需要大量的时间。这解释了经济学中的一个谜团。从启动新域的技术开始建立到新域全面发挥影响，通常需要几十年的时间。电气化的现实技术（电机和发电机）出现于19世纪70年代，但是直到20世纪头20年，它们对工业的全面影响才被人们感受到。詹姆斯·瓦特的蒸汽机发明于18世纪60年代，但是直到19世纪20年代才开始普及。更近发生的数字化技术，如微处理器和阿帕网（互联网的先行者），在20世纪70年代初已进入可使用阶段，但它们的影响，即使在目前数字经济模式已经形成的情况之下，依然没有充分发挥出来。如果你接受我们讲述的关于“采用”的故事，就会知道这些延迟一定是人为造成的，因为人们需要时间去寻找做事的新方式，并确认这么做会改善他们的环境。这样的时间迟延可能会有5年或10年，但绝

不会有30或40年。

一旦我们承认发生的不只是一个简单采用过程，而是一个更大的、发生在域和经济之间的相互采用的过程，这个谜就解开了。对于一场颠覆性改变来说，只有基础技术的改变是不够的。一场颠覆性改变的完全展开需要等我们对那些围绕着新技术的活动（企业或商业流程）进行组织，并且直到这些技术也开始适应我们之后才算真正完成。为了实现这一切，新的域必须积累信徒和声誉，必须找到目的和用途；其核心技术必须能够解决障碍，并且填补组件之间的裂缺；它必须发展它的支撑技术，并且将它和使用它的技术桥接起来；它必须理解它的现象基础以及借由这些而发展起来的理论。

市场必须被发现，现存经济结构必须被重构以便利用新域。旧的配置必须接受新域并熟悉其内在实践方式，这就意味着那些运用旧语法的工程师们要重整旗鼓，面对新域。这样做并不轻松。所有这些必须经由金融、制度、管理、政府政策以及可以熟练运用新域的人共同协作完成。

因此，这一进程需要耗用的时间，不取决于人们从开始注意到不同的做事方式到开始决定采用它的这个过程，而取决于将既有的经济结构进行重构并适应新的域需要花费的时间。它可能需要几十年，而不是几年。在此期间，旧技术会不顾自身的劣势和弱点坚持不懈地存在着。[8]

1990年，经济历史学家保罗·大卫（Paul David）提出过一个可以说明这个过程的经典案例。[9]在工厂还没有实现电气化之前的一个世纪左右，人们依靠蒸汽机提供动力。每一个工厂都有这样一个独立引擎，一个足有几层楼高的巨大的家伙嘶吼着，摇动着一个由活塞、飞轮、滑轮、皮带构成的组合，驱动主轴的运转，然后再继续驱动工厂里其他的机械设备。直到19世纪80年代，电动发动机才作为新的电子域的主要成员进入了可操作阶段。它们所耗能源成本很低，还可以分解成小单元进行安装。这样一来，每个需要提供动力的工作机旁边都可以安装一台电机。同时它们可以做到独立控制，可以依据需要分别开启或关闭。毫无疑问，电动发动机是更优的技术。

那么为什么美国的工厂要花费40年的时间才能采用它们呢？大卫发现，新技术要想被有效地使用，需要不同于旧有的、依照蒸汽机布局结构的工厂建筑。也就是说，新技术要求重新建造工厂。这个代价无疑是昂贵的。此外，即使下决心要重建，但对于工厂需要被重构成什么样子，其实并没有清晰的蓝图。电子工程师们懂得新域的知识，但却不懂

建筑；反过来，工业建筑师懂得建筑，但却不懂电气化技术。因而，将新技术与安置新技术的工业布局设计之间进行整合，并将整合之后的知识传播出去，需要大量的时间。这个案例显示，整个过程用了40多年。仅有企业和人们对新技术的适应还不够，只有新技术自身也开始适应企业和人，才到了真正有所成就的时候。

从某种意义上说，这一结构变化的过程不仅是在经济中单纯耗用时间，它们在经济中实际上是在定义时间。让我来解释一下这是什么意思。传统时间是以标准的时钟时间进行衡量的，但是这里还有另外一种衡量时间的方式，就是通过新结构的“变化”（becoming）来描述时间。哲学家称之为“关系时间”（relational time），意即如果事情总是保持不变，就没有变化来标注事情正在过去，因而也就没有变化来标注“时间”。在这个意义上来说，“时间”将保持静止。同理，如果结构变化了，宇宙间的事物移动并改变它们自己，“时间”就会显现。

在我们的语境下，变化（经济中的时间）的出现是基于经济基本结构自身的改变。这样的情况通常在两个尺度上发生，一个比较快的尺度和一个比较慢的尺度。较快的尺度可以称它为快时间（fast-time），即新技术的设计、测试及被经济吸收的时间，它会显示事物“变化”的步调，以及新的经营活动和新的做事方式的步调。这将按常规时间尺度中的月或年来衡量。较慢的尺度可以称它为慢时间（slow-time），它会出现在新的技术体进入经济领域时，慢时间在经济和社会中显现为时代。它们共同在经济中创造出了“时间”。这将按常规时间尺度中的数年或数十年来衡量。

不论在哪个尺度上，时间都不会创造经济，反而是经济（或经济结构的改变）在创造时间。

创新与国家竞争力

关于新技术体的建构非常引人注目的一件事是，它们的发展前沿通常会高度集中在一个或最多几个国家或地区。纺织技术和蒸汽技术的发展集中在18世纪的英国；一个世纪后的化学技术则在德国得到最大程度的发展；我们这个时代的计算机技术和生物技术则主要是在美国产生和发展的。为什么是这样呢？为什么技术体一定会集中在某个地区而不是平均分散在很多地方呢？

如果技术的力量来自于知识，即关于技术和科学的信息，那么，原

则上讲，任何一个拥有工程师和科学家的国家都应该和其他国家一样具有创新性。毕竟，大多数国家应用的是同样的科学，同样的学术期刊，同样的知识、事实、真理、理念和信息。

但真正前沿的技术，那些处于边缘的复杂技术并不是源于知识，而是源于别的东西，我将它们称为“深奥的手艺”（deep craft）。深奥的手艺不只是知识，它是一套认知体系[10]：知道什么可能发挥作用，什么不可能；知道用什么方法、什么原理更容易成功；知道在给定的技术中用什么参数值；知道和谁对话可以使事情进行到底；知道如何挽救发生的问题；知道该忽略什么、留意什么。这种手艺性认知（craft-knowing）将科学、纯粹知识都视为理所当然。它整体地来自于某种信念的共有文化[11]，蕴含着共同经历的某种无法言明的文化。

这也意味着，它知道如何操纵那些新近发现的还不甚了解的现象，这种认知可能来自当地大学或工业实验室实验性的操作或研究，进而又变成共有文化的一部分。科学，在这个意义上，也是一门手艺。20世纪的前30年，剑桥大学的卡文迪许实验室一直是原子物理学领域发明的聚集地。之所以如此，是因为他们知道如何与原子现象打交道。科学作家布莱恩·卡斯卡特（Brian Cathcart）说过：“无论在这个领域知道些什么，不管是技术、设备、数学工具，甚至理论，它都是被某处的某个人知道的……不仅如此，它还需要接受讨论、挑战和考验，有时是在讨论会上，有时是在其他活动中。对于原子物理学的任何问题或困难，在卡文迪许的某个地方，你一定能找到答案。”[12]

这样的认知根植于地方性微观文化中：在某个公司里、某栋建筑中，在某条走廊上。它们在某处变得非常集中。就这一观点而言，没有什么内容是特别新颖的。阿尔弗雷德·马歇尔（Alfred Marshall）在1890年就曾写道：

> 当一个产业为自身找到这样一个地方，它很有可能就驻留在那里了：人们在邻近的街区学习相同的行业技术，这使他们获得了很大的收益。从此神秘的行业不再神秘。就像呼吸空气一样，孩子在不知不觉中就知道了很多知识。干得好马上会被欣赏。关于设备、工艺以及企业组织形式的发明和改进都会立即被讨论其优劣与否；一个人有了新点子，可能会被别人借鉴过去，并结合他们自己的想法，从而产生了另一个新点子。不久以后，相关的附属产业就会成长壮大起来，从而为主产业提

供工具材料、组织交易，并在很多方面提供有助于构成经济的素材。[13]

自马歇尔时代以来，事情并没有什么改变。如果一定得说有什么区别的话，那就是，行业秘密比以往显得更神秘。这可能是由于它们更多是建立在量子力学、计算科学或分子生物学基础之上。行业秘密或共同认知对于我谈到过的发明、开发以及建构技术体的过程来说，都是完全必要的。所有的建构都需要花时间，且不易转移到其他地方，同时又不可能被完全地记录下来。手艺的形式部分可能最终会成为学术论文和教科书，但真正的专业技能部分则很大程度上藏在它创生的地方，在那里，它被视为理所当然的、共享的，且无须明言。

接下来，一旦一个地区或一个国家因为行业秘密在技术体中领先了，这个地区就会处于更领先的地位。成功会接踵而来，形成对技术的地方性聚集作出的积极反馈或者收益递增效应。一旦一小群公司聚集在新的技术体周围，它就能吸引更多的公司。这就是为什么新的技术体会在一个或两个特殊区域聚集起来，并且很难被挑战。其他地区当然可以为这个新的技术体作出他们自己的贡献，比如参与产品制造或技术改进，但它们不会再有大规模的原创动作了，因为能够提供继续突破所需的详细认知的原产地不在那里。

当然，地区优势无法永远维持。一个地区可以是某些技术体产生的先锋，但是当该地区变得不再那么突出时，它也可能衰落。有时，这种衰落可以通过将一个技术体的专业知识充分嫁接到另一个技术体当中得到遏制。以硅谷著称于世的斯坦福大学及其周边地区，是在1910年以无线电报业起家的，而后在20世纪三四十年代，它将无线电发报技术大量运用到电子工业中去，从而为计算机行业的诞生播下种子，现在它又开始转向生物技术和纳米技术发展了。当新域从旧域中独立出来或者从大学研究中萌芽之后，区位优势就可能建立起来。

技术中不处于领导者地位的国家和地区也并不是毫无希望的。对一些创业公司给予全面、周到的激励，以及投资一些还没定向的基础科学，都会很有帮助。因为技术总会在不经意间播撒许多带有活性的种子，因此如果种子恰好落到恰当的地方，某个集群过程就可能在某个意想不到的地方产生。

在20世纪80年代，俄亥俄州阿克伦地区的轮胎行业正面临着全球竞争、产品召回事件的困扰。[14]几家大型轮胎公司，如B.F.古德里奇、凡

世通轮胎公司和通用轮胎公司等，都已经撤离该地区。但阿克伦从开创它的橡胶时代之日起就拥有强大的高分子化学（化学分子链）专业知识，它设法将这种优势应用到更广泛的一套认知体系中去，那就是高技术聚合物制造。结果是，阿克伦现在已经成为名副其实的“高分子谷”（Polymer Valley），拥有400家从事相关业务的公司。对一项技术的深层认知可以被利用到另一项技术的深层认知中。

所有这些都会造成国家间的竞争。[15]技术的发生始于对现象的深入理解，而这将逐渐内嵌为一套寓存于人的、地方性自我建构的、深邃的共同认知（shared knowings），并将随时间而发展。这就是在科学上领先的国家在技术上也会处于领先的原因。因此，如果一个国家希望能够引领先进技术，它需要的不是投资更多的工业园区或含糊地培养所谓“创新”，它需要建立其基础科学，而且不带有任何商业目的。它应该在稳定的资金和激励安排下养育那样的科学，让科学在一些初创的小公司中自己实现商业性的发现，并受到最少的干扰，要允许这些新生的冒险者成长、萌发，允许这门科学及其商业应用播种新的颠覆性改变。

这并不是个很容易就能自上而下得到控制的过程。对政府来说，总有一种诱惑让科学去追逐某个特定的商业目标。但事实上，这样很难奏效。如果20世纪20年代的量子物理学曾宣称过要达到什么商业目的，那它一定是以失败告终的。但是，量子物理学带给我们的是晶体管、激光、纳米技术的认知基础以及更多其他东西。产生先进技术的能力和社会主义的计划生产体制不同，它更像从事某种园艺，需要种植、浇水、除草，这比弄个什么几年计划更有效。

如我在本章所述，应该清楚的是，域的发展和个体技术的发展是不同的。这个过程和喷气式发动机开发中那种专著的、聚精会神的、理性的过程不太一样，它更像是一个系统的法律条款的形成，是一种缓慢的、有机的、累积式的过程。对于域来说，产生出来的不是一种新设备或新方法，而是一个新的表达方式的词汇表——一种为产生新功能而进行编程的新语言。这个过程也是缓慢的。围绕着对一系列现象的松散理解或一种可行的新技术，一个域会逐渐成形，并且有机地建构在支持这些元素的组件、实践和理解的基础之上。当新域到来时，经济会遭遇它并最终改变自身。

而所有这一切都是创新的另一个切面。[16]事实上，我们可以将前面4章看成是关于创新的详细解释。这里没有一种单一的机制，而是大约4个相互独立的机制。创新存在于新的解决方案转变为标准工程的过程

中，其间包含着许许多多微小的进步和修正，它们累积在一起共同推动着实践前进；创新存在于由发明引发的根本性新技术产生的过程中；创新存在于这些新技术在改变内部组件或者结构深化时，因增加组件而获得发展的过程中；创新还存在于技术体从出现到随时间而发展，最后创造性地改变了那些与之遭遇的产业的过程中。

这里谈的每类创新都很重要。并且它们中的每一种都摸得着、看得见，创新不是什么神秘的事。它的发生不是模糊地求助于所谓“创造性”。创新实际上只是另辟蹊径地完成经济的任务。

面对我研究的案例，我一次又一次强烈地意识到，创新总是出现在人们面临问题的时候，尤其是在面对那些非常清晰的问题时。创新总是作为解决问题的方案出现，这些方案是由那些对组合手段或功能着迷的人想出来的。创新的增强可以通过资助创新必须的支撑因素，通过在无数的功能中经受磨炼并培养经验，通过建立专项研究和实验室建设，以及通过当地文化滋养深层认知来实现。但创新不会为某个地区、某个国家或某个人所垄断。它可以发生在任何地方，只要那里的问题能够被研究，并具备形成解决方案的足够多的背景资料。

事实上，创新有两个主要的主题。一个是如何不断在现有工具箱里的零部件及实践中去发现或组合新的解决方案；另一个是产业如何不断将它们的实践过程同那些来自新的工具箱（新域）的功能组合起来。第二个主题与第一个类似，也是关于新的过程和制度安排的创造，是实现目的的新手段，但是它更为重要。这是因为重要的新领域（例如，数字领域）遭遇到的是经济中的所有产业。当这一切发生的时候，域将它们所能提供的解决方案与许多工业领域内部原有的制度安排方式结合起来。结果是，新工艺、制度安排、做事的新方式不仅仅是在一个地区内被使用，而是将贯穿到整个经济中。

最后一个评论。在本章和第7章，我都谈论了技术“发展”的问题：当单个技术或技术体成熟的时候，它们都会进入一个可预测的阶段。我本来也可以说它们是在“进化”。每个技术或技术体都有它的后代，所有的分支又会有不同的“亚种”或不同的次级域参与进来，从这个意义上说，它们的确是在进化。但是我还是使用了“发展”这个词，因为我宁愿将“进化”这个词留给整个技术（对一个社会有用的人工物和方法的集合），描述它是如何从那些已经存在的元素中创造出新的元素，并以此为基础进行扩建。

这将是本书的中心议题。我们已经收集了我们所需的全部材料，那

么我们现在就开始吧。

THE

NATURE

OF

TECHNOLOGY

第9章 进化机制

组合是新技术的潜在来源。组合的威力，在于它的指数级增长。如果新技术会带来更多的新技术，那么一旦元素数目超过一定阈值，可能的组合数就会爆炸性增长。此外，机会利基也在呼唤着新技术。技术就如同生命体一样，它的进化与生物进化也没什么本质差异。

想象一下，将人类所有的技术都收集在一起会是什么样子。把过去的、现在的、所有曾经存在的各类技术，所有的工艺、设备、部件、模块、组织形式、方法，以及正在应用的或者过去曾经应用过的计算方法都堆在一起，那会是个什么景象？如果我们真能为此列一个目录，那么它的数量将是巨大的。

这是一个技术的集合，现在我们想要探究它是如何进化的。

我一直声称这个集合的进化是个自我创造的过程：新元素（新技术）的构成来自于那些已经存在的元素，而这些新元素又能为进一步的建构提供建构模块。现在我想要弄清楚这一切发生的机制。

窥一斑可见全豹，你可以从这个集合的一个角落去观察这个技术自创造过程的微缩图景。在20世纪初期，李·德·福雷斯特（Lee de Forest）[1]一直在试验改进无线电信号检测手段，他尝试在一个二极管中再插进第三个电极，他期待他的三极管能使信号放大。因为当时无线电信号传输能力很弱，所以这样的诉求在当时的情况下显得非常强烈。但是他没有成功。在1911年和1912年期间，包括福雷斯特自己在内的一些工程师几乎同时致力于如何设法把三极管与已有的电路元件合并起来以生产出一种可行的放大器的工作。放大器电路与标准元件（线圈、电容器及电阻）进行了一些稍有不同的组合之后便产生了振荡器电路，这种电路能产生出被当时社会寄予厚望的东西：纯粹的单频无线电波。当然，它依然需要其他标准元件的配合才能使现代无线电传输机与接收机成为可能。之后还要再通过结合其他元件才能使无线电广播最终成为可能。

而这还不是故事的全部。在一个稍有不同的电路中，三极管还可以被用作继电器：它能作为一个开关，通过控制三极管栅极的电压来开合电路。如果断开开关，继电器可以表示为0或逻辑值为“假”；如果连接开关，则表示为1或“真”。将继电器进行恰当地组合就可能产生原始的逻辑电路。这样的逻辑电路与其他逻辑电路和电子元件一起工作，就使早期计算机成为可能。因此，在其后大约40年的时间里，无论对无线电还是对现代计算技术来说，三极管都是一个承前启后的关键要素。

技术自身创造了自身。它从已有的技术集合中一点一点建构起来。我这里想要做的是详细描述这一切是如何发生的，即技术是如何进化的，它是如何从最初如此简单的形态发展到现在这样一个不同寻常的复杂世界的。

我一直在讲技术是从已有技术中被创造出来的[2]（或者是从已经存在的技术中被创造出来的），让我解释一下为什么事实如此。任何解决人类需求的方案，任何达到目的的新手段，都只有通过使用已有的方法和组件才能使其在现实中实现。因此，新技术的形成（或成为可能）总是源于现有的技术，而且总是如此。如果没有压缩机和燃气涡轮机，没有机床来制造这些高精度的机器，喷气式发动机是不可能产生的。聚合酶链式反应是将分离DNA、解旋、附上引子、重建双链DNA等方法组合起来完成的。这是一种对已经存在的事物的组合。

读者可能会反对说，应该也有例外，比如青霉素。作为一种治疗的手段，青霉素是技术，但它似乎并不是以往技术的组合。但请思考一下：把弗莱明发现的基本现象转变为一个可行的治疗手段是需要一套非常明确的已有技术的。它需要生化过程去隔离霉菌中的活性成分，需要其他程序对其进行纯化，还需要生产和传送等过程。青霉素技术在这些手段和方法方面都有它的母系或来处，它不可能在一个不具备这些支撑元素的族群中产生。所以，是现有的手段使青霉素技术成为可能的。所有技术产生于已有技术，也就是说，已有技术的组合使新技术成为可能。

当然，造就一个使新技术成为可能的元素可能不仅需要单纯的物理条件，还需要其他包括生产或装配等在内的条件。确证技术的亲子关系也并不容易：将青霉素带到世界来的技术和方法有许多种，那么，哪一个才算是它的父母呢？我们可以说一定是那些重要的技术和方法，但当你想确定哪些是而哪些不是时，在某种程度上只能凭感觉行事了。但是，这并不会干扰我的核心观点：所有技术的产生（或成为可能）都源自以前的技术。

这又将我们置于何处了？严格来讲，我们应该这样说，新元素形成的可能性是由既有元素促成的。但如果说得宽泛些，那就是，技术是从已有的技术中产生的，是通过组合已有技术而来的。正是在这个意义上，技术集合中新元素的产生（或成为可能）源于已有的技术集合，结果就是技术创生于技术自身。

当然，说技术创造了自身并不意味着技术是有意识的，或能以某种阴险的方式利用人类为它们自身的目的服务。技术集合通过人类发明家这个中介实现自身建构，很像珊瑚礁通过微小生物自己建构自己一样。所以，假如我们把人类活动总括为一类，并把它看成是给定的，我们就可以说，技术体是自我创生的，它从自身生产出新技术。或者，我们可

以采用温贝托·马图拉纳和弗朗西斯科·瓦雷拉自创的一个词——自创生，来描述这种自创生系统，这样就可以说技术是自创生的（autopoiesis）。这个词的希腊语是“自我创造”（self-creating）或“自身涌现”（self-bringing-forth）的意思。

自创生看起来有些抽象，更像某种系统的哲学理论。但事实上，它告诉了我们更多东西。它告诉我们，每一个新技术都是从已有的技术中来的，因此每项技术都站在一座金字塔之上，而这座金字塔又是由别的技术在更早的技术之上建立的，这个连续的过程可追溯到最早人类捕获的现象。它告诉我们所有未来的新技术都将来自现存技术（也许是以一种不明显的方式），因为它们都是构成未来新元素的元素，而这些新元素将最终使未来新技术成为可能。它告诉我们，历史是重要的：如果技术由于某种偶然以不同的序列出现，建立在这些技术之上的技术也会不同。技术是历史的产物。它还告诉我们，技术的价值不仅在于可以用它做什么，而且在于它进一步可能导致什么。技术专家安迪·格罗夫（Andy　Grove）曾经被问：网络商业的投资回报是什么？“这好比是在问正在注视着新世界的哥伦布”，他答道，“什么是他的投资回报呢？”

自创生给人一种感觉：技术是通过扩展延伸到未来的，也给了我们一种去思考人类历史中的技术的方式。通常，历史呈现的是一套发明，它们发生在不同的时间，并且是不连续的，有时会有一些交互的影响。如果我们用这个自创生的观点重新回溯技术的起源方式的话，这个历史看起来会像什么？让我在这里作一个简略的回顾。

最初，第一个被利用的现象是自然界能直接呈现的。[3]有些材料被切削时出现了薄片状，这就是燧石或黑曜石带刃工具的根源；重物在坚硬的表面上可以碾碎东西，这是碾磨草药和种子的根源；柔韧材料通过弯曲可以蓄积能量，这是用鹿角或小树叉制成弓的根源。这些现象自然地直陈在大地上，使原始的工具和技术成为可能。接下来，这些原始工具技术再继续使其他技术成为可能。火使烹饪成为可能；挖空的原木使最原始的独木舟成为可能；烧制使陶器成为可能。由此又开启了更多其他的现象：某些矿石在高温下产生可以被塑形的金属，又因之产生了武器、凿子、锄头以及钉子。元素的组合就此开始发生：用纤维编织的带子或绳索被用来将木柄和金属绑到一块，这样就组成了斧子。技术和工艺实践的集群开始出现了，如染色、制陶、纺织、采矿、打铁、造船。风和水被用作能源。滑轮、杠杆、曲柄、绳索、齿轮的组合出现了，这些早期的机器被用于碾磨谷物、灌溉、建筑、计时。围绕在这些技术周

围的工艺实践随之也有了进步。还有通过实验过程获取的对一些现象及其应用的粗浅理解。

随着时间的流逝，这些理解提供了近距离观察现象的方法，随后对这些现象的利用开始系统化起来，也就是开始使用科学的方法——这是现代纪元的开始。化学、光学、热力学、电学的现象开始被理解和应用，通过仪器的使用（例如，温度计、热量计、扭秤）达到了精确的观察水平。技术的巨大领域开始运作，如热机、工业化学、电力、电子等。借由这些，人们又发现了更进一步的现象：X射线、无线电波传输、相干光等。进而是激光、无线电传输以及逻辑电路元件的不同组合——现代通信和计算就此诞生了。

就是这样，由少及多，数量多了就成了专业的，专业化以后再发现更多的现象，使更好地利用自然法则进一步成为可能。到了现在，随着纳米技术的来临，被捕获的现象可以自己直接再去捕获现象，通过去移动和放置材料中的单个原子来达到更进一步的特定目的。所有这一切都起步于地球最初的自然现象。如果我们当初居住的世界具有完全不同的现象，我们就会拥有不同的技术。照此看来，从人类的角度来衡量，这是个漫长的过程，但从进化的角度来衡量，这却是一个短暂的过程。技术体就这样被它自身建立、深化、专业化、复杂化了。

我在这里说得简略，我没有谈到这一切发生的机制。在这章的剩余部分，我要去详细阐述技术进化的实际机制。这些将是我的理论的最核心部分。

请允许我先谈谈主导技术发展的更大力量，然后再把镜头推向更具体、更细节的机制。其中一股力量当然是组合，我们可以将它视为现有的技术体“供给”新技术的能力，无论是将现有的零部件进行总成，还是用它们捕获现象。另一股力量是需求，对要完成目的的手段的需求，还有对新技术的需求。将这些供给和需求整合在一起，就产生了新元素。让我们对比逐一加以讨论。

组合

关于组合，我已经说了很多。但是，作为新技术诞生的潜在来源，它到底有多强大呢？

当然，我们可以说，随着技术数量的增加，组合的可能性也相应提

高。事实上，威廉·奥格本早在1922年就已观察到这一点：“可用于辅助发明的东西越多，发明的数量就越大。”[4]事实上，他预期物质文化（技术）的增长显示出“和复利曲线的相似之处”。如果他能在今天进行写作，他会说技术是按指数级增长的。

奥格本对他的指数推断没有给出理论上的支持，但我们可以提供一些，这只需要一些简单的逻辑。假设一个技术集只包括技术A、B、C、D及E。新的可行的组合可能包括不同结构的模块（例如，AED或者ADE）。并且它们可能不止一次包含同一技术：其中可能有冗余（例如，ADDE和ADEE）。但是让我们保守一点，只纳入包括或没包括建构模块的可能性。即，没有不同的体系结构，也没有冗余。这就可以使我们有诸如AB、AE、BD这样的双重组合的可能性；或者像CDE、ABE那样的三重组合；以及类似BCDE、ACDE那样的四重组合等。

到底有多少这样不同类型的组合？在给定的新组合中，每种技术，A、B、C、D、E可能出现或不出现。这样就有A或B或C或D或E出席或缺席——对于A（出席或缺席）有两种可能，然后是B出席或缺席。从A到E计数之后，就是二乘以二乘以二乘以二乘以二，即2^5或者32种可能性。我们需要减去仅有一个技术出席的情况，即仅有A或B或C出席（这些不是组合），以及0个技术的情况，即没有建构模块呈现。我们现在用32减去原初的5种情况，再减去0技术的情况，余下26种可能性。一般而言，对于N种可能的基本元素，我们会得到$2^N–N–1$种可能的技术。对于10个模块元素，我们有1 013种组合的可能，20个则有1 048 555种可能，30个则有1 073 741 793种可能，40则有1 099 511 627 735种可能。这种可能的组合以指数方式变化（例如以2的N次方变化）。对于任何特定数量的模块，组合的可能性是有限的，在数量很少的情况下，它们看起来并不很大。但是一旦超出这些小数量，它们就立即变得极为庞大起来。

当然，并不是所有的组合对工程都有意义。GPS的芯片技术可能需要很多可能性，而与喷气式发动机相比，鸡窝需要的可能性则可谓少之又少。但只要模块可以由一套既定的模块制造出来，即使是百万分之一的机会也是起作用的，其可能性依然是以（$2^N–N–1$）／ 1 000 000或大约是2^{N-20}。也就是说，它们还是以指数模式增长的。

必须承认，这样的计算是粗略的，但是我们可以用几种方式来加以改善。我们可以认为许多组合不具经济学意义，因为相对于要实现的目的而言，它们过于昂贵。还有一些，如激光或蒸汽机，可能引发一连串

进一步的装置和方法，而另一些可能什么也留不下。我们可以允许同样的组件在不同的结构中多次使用。仅仅一个电子部件，如晶体管和其自身的复制品构成的组合就可以创造出巨量的逻辑电路。细微的改良有很多，但我想提请读者注意的，也是我尝试表明的观点是，即使是这些粗略的组合论，它们也表明了：如果新的技术会带来更多的新技术，那么一旦元素的数目超过了一定的阈值，可能的组合机会的数量就会爆炸性地增长。用相对较少的模块进行组合，就会有无限的组合。

> 技术思想前沿
>
> 技术的进化机制就是“组合进化”。所有技术都是从已经存在的技术中被创造出来的。如果新的技术会带来更多的新技术，那么一旦元素的数目超过了一定的阈值，可能的组合机会的数量就会爆炸性地增长。有些技术甚至以指数模式增长。

机会利基

即使已有技术的组合为新技术提供了一件潜在的“供应”，它们也只会因出现了某种需要，或说某种“需求”而产生。事实上，需求不是一个恰当的词。因为在青霉素或核磁共振还不存在的时候，根本无法在经济中对它们提出明确的需求。因此不如说我们应该谈论的是技术的机会——即技术可以有效地占据的利基。机会利基（opportunity niches）的出现召唤新技术的诞生。

那么在人类社会或经济中，究竟是什么导致了机会利基的产生？

显而易见的答案是，人类的需求产生了社会利基。我们需要被遮挡、被喂养、迁移、保持健康、穿衣以及娱乐。有一种倾向认为，这些需求是固定的，就像有一个带有子目录的需求清单。但是当你深入其中任何一个类别的需求时，比如，住的需求，你会发现它不是固定的，它将在很大程度上取决于社会发展的状态。我们想要的庇护场所（住房）在很大程度上取决于什么人住在什么里面，什么人拥有什么，什么人标榜什么，这就如同在瞥了一眼《建筑文摘》之后需要作出确认一样。当基本的需求被满足，社会达到一定程度的繁荣之后，这些“需求”开始像血管扩张那样进行分化。对娱乐的“需求”在早期的社会是通过公开的展示或讲故事来完成，而现在则需要全套的体育运动、戏剧、电影、小

说、音乐。在这些基本需求得到满足后，就会又有新的分支繁衍出来——例如，我们有许多种类型的音乐。

对机会利基产生的原因来说，应该还有别的答案。人的需求不仅是由技术带来的社会繁荣创造的，而且还直接来自于技术自身。一旦我们有办法诊断糖尿病，我们就产生了一种人类需求，一个机会利基——一种对控制糖尿病的手段的需求。一旦我们拥有火箭技术，我们就会产生一种对太空探险的需求。

和许多人类生活的其他方面一样，我们的需求是精致细腻的：它们对社会状况的依赖既精妙、轻松，又复杂、异常。它们会随着社会的繁荣而愈加复杂。由于是技术促进了社会的繁荣，所以也正是技术促进了需求的成长。

技术思想前沿

机会利基的出现召唤新技术的诞生，绝大多数机会利基的产生缘于技术自身，这是因为以下三个原因：

- 每个技术通过它的存在建立了一个能够更经济或更有效地实现其目的的机会。
- 每项技术总是需要另外的支撑技术来制造它，这些支撑技术又需要它们自己的次级支撑技术。
- 技术经常引发间接性的问题，这会产生需要提供解决方案的需求或者机会。

但这还远远不是答案的全貌。绝大多数技术创造的利基市场不是缘于人类的需求，而是缘于技术自身。之所以如此，原因很多：其中一个原因是，每个技术通过它的存在建立了一个能够更经济或更有效地实现其目的的机会；另一个原因，每项技术的存在总是需要另外的支撑技术来制造它，组织它的生产、分配、供给，提高它的性能等，而这些支撑技术反过来又需要它们自己的次级支撑技术。

汽车在1900年创造了一套与之配套的需求（机会利基市场）：需要流水线生产方式，需要铺设道路和汽油提纯，需要维修设施和加油站。接下来，汽油又产生了进一步的需求：炼油厂、原油进口以及石油的开采。

技术之所以能够催生机会利基，还有第三个原因。技术经常引发间接性的问题，这会产生需要提供解决方案的需求或者机会。在17世纪的时候，欧洲比较容易开采的矿山储备都被开发殆尽了，进一步的开采需要深层采掘技术。渗水此时成为一个很大的问题，因此需要更有效的排水手段。这一问题在1698年由托马斯·萨弗里（Thomas Savery）发明的一款原始版的蒸汽机解决了（尽管不太成功），他号称这是一项“以火力推进的，用于排水以及所有类型工厂所需动力的新发明”。

对于我们的论题而言，我们不需要一个关于人类和技术需要如何形成的完整的理论。但是我们却必须意识到，这个系统不仅包括技术创造技术，也包括技术创造的机会利基对技术的召唤。我们还需要意识到，技术的机会利基不是固定和既定的，而是在非常大程度上由技术自身产生的。机会利基市场随着技术体的变化而变化；它们随着技术体的生长而生长，并逐渐复杂起来。

核心机制

这些驱动的力量给我们提供了一个宏大的图景，已有技术的组合提供了新技术的可能性：一种潜在的供应。而人类和技术的需要又创造出了无数的机会利基市场：一种需求。随着新技术的出现，需要进一步驾驭和组合的新机会持续出现着。所有这一切都自展式地前行。

那么这种自展是通过什么具体的步骤、机制来实现的呢？

可以将技术体看作是一个网络，这个网络是自我建构的，并且有机地向外部生长。在这个网络当中，每个技术（我会将它称为元素）都表现为一个点或节点。每个节点都和指向它的母节点相连，正是这些母节点使这些新节点成为可能。当然，在给定的时间之内，并不是所有的技术都会被积极地应用到经济当中。可以设想有些节点或元素是蓄势待发的，我将这些元素称为“活跃技术体”：这些元素在经济上是可行的，并被应用在当下的技术中。另一些元素，如水车以及帆船，则在本质上已经死了，它们从活跃技术体中消失了。它们有可能在新的组合中被重新启用，但是这种情况很少发生。

新技术不时地加入到活跃技术体中，但不是整体性加入。活跃的网络在一个时间点可以在某地迅猛发展，而在另一些地方却完全不行。有一些元素通常伴随着最近被捕获的现象（例如，20世纪60年代的激光），它们会迅速生产出进一步的元素；而另一些成熟并已建立起来的

元素，如索尔维制碱法，则已经没有后代诞生了。活跃网络的建构是不均衡的。

随着活跃技术体中元素的增减，利基市场的集合也发生变化。我们可以将这些需求想象成被发布在一个巨大的布告栏中。（我们可以认为工程师和企业家正在观看这个布告栏并对此作出反应。）每个新元素都必须满足布告栏上至少一个需求或目的。随着新元素加入到网络中，那些曾经满足过目的的旧元素，或者那些不再具备经济性的元素会脱离网络，它们的机会利基市场也会在布告栏上消失。调节所有这些的力量是经济。我们可以把经济看作是这样一种体系：它决定成本和价格，并因此标注需要新元素完成的机会，同时决定哪位候选人可以进入活跃技术体。（到目前为止，我还是会将经济作为给定的黑箱来对待。对此我将在第10章进行更多的讨论。）

让我们探寻一下技术的活跃网络是如何建立起来的。它通过长期以来借由一系列新技术的可能性，并与当时的机会利基遭遇而发生了进化。它们的每次遭遇都是工程的和经济的。候选的解决方案对于要达成的目的而言，必须在技术上“可行”，才会被予以考虑，而其成本则必须符合市场支付意愿。符合这些条件的技术是达到当时目的的潜在的“解决方案”。这样的解决方案可能同时有几个，最终加入活跃技术体的元素就来自其中。

我们最初的认识更多地不是来自技术的稳定聚集，而是新元素和机会利基的形成过程，以及它们的变迁和消失。这个过程是可计算的：它是按照离散步骤进行的。我们可以从假设一个备选新技术的出现开始。它已经借由先前的技术的组合成为可能，并且已经击败了竞争对手进入到了经济领域。接下来会有6件事情或6个步骤发生。我们可以把它们看作这个技术建构游戏中的法定步骤。我可以对此进行抽象的阐述，但读者可能发现，如果这时心中有一个现成的技术案例，将会对理解它们很有帮助，比如说，晶体管。

1．新技术作为新元素进入到活跃技术体当中，它就变成了活跃技术体中的一个新节点。

2．新元素可能取代现有技术或现有技术的零部件。

3．新元素为支撑技术和制度安排建立进一步的“需求”或机会利基。

4．如果旧的、被替换了的技术逐渐退出技术体，它们的附件也要被丢弃。随其而来的机会利基也将和它们一起消失，填补了机会利基的元素也可能就此不再活跃。

5．作为未来技术或未来元素的潜在元器件的新元素将活跃起来。

6．社会经济（商品和服务的生产和消费模式）会进行重新调整以适应这些步骤。成本和价格（也因此成为刺激新技术产生的诱因）也会作出相应的变化。

就这样，晶体管在20世纪50年代进入到了活跃技术体当中（步骤1）；在众多的申请者中脱颖而出取代了真空管（步骤2）；建立制造半导体设备的需求（步骤3）；导致了真空管工业萎缩（步骤4）；成为许多电子设备的主要组成部分（步骤5）；迫使电子设备的价格和诱因作出改变（步骤6）。

但这样罗列下来，事情看起来太前后有致了。在实践中，它们其实并不能这样干净利落地前后相继、井井有条，而经常是并行发生的。还是以晶体管为例，一个新的技术变成一个潜在的模块（步骤5），在它一出现就形成了（步骤1）。而新机会（步骤3）的出现也和新技术的出现如影随形（步骤1）。当然，这些步骤当中的任何一个都需要耗用时间来完成。技术通过经济进行传播需要时间，而经济反过来又需要很长时间来适应新的技术。

如果这些步骤都是按次序在每个时段发生一步，技术体的建构过程就有其方法了。每一个新的可能性将添加一个元素，然后其他5个步骤就会及时跟进。但一些更有趣的事发生了。每个步骤都可能勾动一连串的其他步骤。一种新技术可能会导致对该技术体的一系列进一步的增补（通过步骤3和步骤5）。价格的调整（步骤6）可能造成某个备选技术突然变得切实可行并进入到活跃进程中去（步骤1）。所以，这些步骤自身就可能引起新一轮的对技术体的添加。每出现一项新技术都可能启动新的事件，且永无尽头。

还有另外一种可能性。一种新元素不仅可以引起它所替换的技术的崩溃（步骤2），而且还能引起依赖被替换技术的需求的那些技术的崩溃（步骤4）。随着这些次级元素被替换，它们的从属机会利基也就崩溃了，占据这些从属机会利基的技术也一起崩溃。20世纪早期，汽车的出现引起了对马车运输的替换。马车运输的死亡又引发了铁匠和马车制造的消逝。铁匠的消逝反过来又导致了铁砧制造的消逝。崩溃呈现一种

逆向链式的规律。这不太像熊彼特所说的“破坏性创造风暴”（gale of creative destruction）[5]，即新技术会消灭广泛存在于经济中的某种商业和工业。相反，它是一个多米诺骨牌式的连锁崩溃，或可称之为“雪崩式的毁灭”。

具有创造性的一面是，正如熊彼特指出的那样，新技术和新工业取代了那些崩溃了的技术或工业。我们可以补充的是，新技术可以轻易地建立起新的机会利基，去静待更新的技术来占据它，而更新的技术又会建立更新的机会利基，更新的机会利基再一次等待更更新的技术来占据。这里也有机会创造的雪崩，也许我们应该称之为风潮（**winds**）。

所有这些活动同时在网络的许多节点上发生着。如同生物圈物种的建构一样，这是一个并行过程，而且毫无规则可言。

我刚才所谈的是关于技术进化的一种抽象的步骤，一种算法。如果我们以几个原始技术开始，然后在头脑中想象这个运动着的系统，我们会看到什么呢？我们是否能看到任何类似我早前所作的关于技术进化历史的那种描述？

好吧，如果我们想让运算法则起作用的话，首先，过程是缓慢的。不仅技术是稀少的，机会也是稀少的。曾几何时，达成一个目的只需要通过驯服某些简单的现象就可以实现，比如在我们人类历史中对火的应用，或者利用藤条进行捆绑。但是这些原始技术提供了机会，至少提供了进一步改良的机会。当机会匹配了，其他的原始技术就会出现，也有可能替换已有技术。技术的积累建立起来了，可用的模块也随之积累起来。机会利基的囤积同样建构起来了。新的组合或技术融合开始成为可能。当新的模块形成，进一步组合的机会就增加了。此时建构过程就变得忙碌起来。组合开始成型于组合，这样一来，原本简单的东西现在变得复杂起来。组合替换了其他组合中的零部件。机会利基开始多元化起来。随着新组合创造更新的组合，系统中的元素出现了爆炸性增长。而随着被替换组合连同其支持技术的机会利基一起被替换，解构的雪崩就开始了，这又引起满足这些机会利基的技术以及支持这些技术的机会利基的进一步消失。这种坍塌因规模和持续时间不同而不同：有一些坍塌的规模很大，大多数则很小。从整体来看，技术集合的总规模在增长。但是活跃技术体部分也各具规模，并且如我们期望的那样，展现了一个随时间增长的网络。

这种进化一旦开始就没有任何理由结束。

关于进化的一个思想实验

我已给出的图景是想象出来的，它是一个关于技术进化的机制如何运作的思想实验。[6]假如我们能以某种方式在实验室中或者计算机上实现这些步骤的话就更好了——比弄出某些技术进化模型来要好。这样一来，我们就可以在现实中观察这种进化了。

但要实现这种设想很困难。技术在形态上大相径庭，而且识别某些组合（例如，造纸和哈伯制氨法）是否有用或有什么用，对计算机来说非常困难。但是我们可以把自己限定在某个技术世界之内，它可以在计算机上发生进化，这样一来，我们就可以对其进行研究了。

最近我和我的同事沃尔夫冈·波拉克（Wolfgang Polak）建立了一个在计算机上展现的人工世界。在我们的模型世界中，技术是逻辑电路。某些读者可能不熟悉这些，那就让我先解释一下。

我们把逻辑电路看作带有输入和输出针的微电子芯片。输入给定电路是二进制形式的数字1或者0，或者也可能是真和假的某种组合，代表那些目前实现的情形。这样输入到一个飞机里的逻辑电路中的输入信号可以检查该引擎在A、C、D、H、K哪种情况下为真或为假，以此表明某种状态，比如说燃料条件、温度或压力；输出针则发出信号，指示开关Z、T、W和R应该“开”（真）或“关”（假），以此来控制发动机。电路会因目的不同而不同，但对于每一组输入值，给定电路会在输出针安排一套特殊的输出值。有趣的是，计算电路和运算法则中的操作是相符合的，例如和：输出值是所有输入值的加和；或者它们是符合逻辑操作的，例如，3-bit AND（即如果所有输入针1、2和3，都为真，那么输出针信号则为真，否则则为假）。

研究逻辑电路的工作对我和波拉克有两个好处。我们始终知道逻辑电路的准确功能。如果我们知道一个逻辑电路是如何连接的，我们（或计算机）就可以准确找出它究竟做了什么。如果计算机将两个逻辑电路组合在一起，这样一个电路的输出就成为另一个电路的输入——这就给了我们另一个逻辑电路，而我们依然知道它明确的功能。因而，我们总是知道组合的功能以及它们的有效性。

我和波拉克将想象我们的人工世界住满了人，有小逻辑学家和小会计师，他们在这个逻辑的世界里急于统计、比较事物。一开始，他们没有工具或手段，但是他们有长长的需要某些特定逻辑功能的需求清单。

他们可能需要能运行AND，Exclusive-ORs，3-bit　　addition，4-bit EQUALS等指令的电路。（为了简单起见，我们假定这个需求清单或者机会利基清单是不变的。）我们计算机实验的目的是要知道技术系统（逻辑电路）能否进化，即能否通过组合已有的元素去填充清单上的机会利基，并在进化发生时对它进行研究。

正如我所说的那样，在实验开始的时候，没有机会利基得到满足。所有可用的技术只有NAND（不是AND）电路（可将此看作是原始的电路元素，一个不会比几个晶体管更复杂的电脑芯片）。接着在实验的每个步骤中，新的电路会通过组合（随机地连接电线）被创造出来。（开始的时候，它们依旧只是NAND电路）。当然大多数新的随机组合都将和需求失之交臂，但是总会有一次组合可能偶然地与清单上的需求相匹配。这种组合一旦出现，计算机马上得到指令将此作为一个新技术或新的模块元素纳入其中。此后它就成为进一步连接和组合电路的有效模块元素了。

波拉克计算机里的这个技术进化实验是自主运行的存在，一旦我们按回车键启动实验后就没有任何人为介入了。当然它可以被一遍一遍地重新启动，以此来比较每一次重启之后有什么不同。

我们发现了什么呢？开始时，只有NAND技术存在，但是在成百上千步组合之后，满足简单需求的逻辑电路开始出现了，并且成为进行进一步组合的建构模块元素。在应用这些模块进行组合之后，能满足比较复杂的需求的技术出现了。在大约25万步组合（20个小时的电脑运行时间）之后，我们停止了它的进化并开始检查结果。

我们发现，在足够长的时间之后，系统已经进化成极为复杂的电路：在其他一堆方程式中，还有8-way-Exclusive　OR，8-way　AND，4-bit　Equals。在运行几轮后，系统进化出一个8-bit加法器，这是一个简易计算器的基础组成部分。这看起来好像没什么特别了不起的，但实际上它是惊人的。一个8-bit加法器拥有16个输入针和9个输出针。如果你再次运行一些简单的组合数学，结果是超过10的177　554次方的可能性会产生16输入针和9输出针的电路，但是其中只有一个可以正确运算。10的177　554次方是一个非常巨大的数字。它远远大于宇宙中基本粒子的数量。事实上，如果我想把这些数字写下来，它可能会占去这本书一半的页数。因此，这样一个电路被250　000步随机组合发现的机会是可以忽略不计的。如果不知道这种进化的运行过程，在实验结束打开计算机屏幕时，当你发现电路已经克服那么长时间的试错从而进化出一个能

正确运算的8-bit加法器，你一定会惊讶于出现了这么复杂的东西。你可能会猜想机器里住着一位天才的设计师。

我们这个过程之所以可以产生这样复杂程度的电路，是因为它首先创造了一系列“垫脚石技术”（stepping-stone technology）。它能够创造满足简单需求的电路，以这些电路为模块，再创造中等复杂程度的电路。然后再用它们创造出更复杂的电路，以自展的方式解决复杂需求。更复杂的电路只有在那些简单电路就位之后才能被创造出来。因为我们发现，当我们取消聚集了垫脚石技术的中等复杂的需求后，复杂需求将无法得到满足。

这表明，在现实世界中，如果没有无线电或对无线通讯的需求，雷达可能不会得到发展。生物界也是如此。[7]如果没有中间层级的机构（例如，从黑暗中辨别光的能力）以及相应的“需求”（从黑暗中辨别光的能力的有用性），诸如人眼等复杂的生物特性是不会出现的。

此外，我们还另有斩获。当我们检视进化的历史细节时，我们发现了一个巨大的时间缺口，在这个缺口里几乎什么都没有发生。然后我们发现一个关键电路（一个新技术）会突然出现，并且迅速将它用于更进一步的技术。一个完整的加法器电路可能需要32 000步才能出现；而之后2-bit、3-bit、4-bit加法器的出现则迅速得多了。换句话说，我们发现了休眠周期，以及随后的缩微版“寒武纪大爆炸”式的快速进化。

我们发现，进化是具有历史依赖性的。实验中每次运行都会出现相同的简单技术，但是出现的顺序则不同。因为更复杂的技术是从简单技术中建构出来的，所以它们常常组合于不同的模块。（如果铜在铁之前出现，许多人工物就会用铜制造；如果铁比铜早出现，则同样的人工物会用铁来制造。）我们还发现，某些对电路的复杂需求则根本得不到满足，比如带有许多输入针（每次都不一样）的加法器或者比较器便是如此。

此外，我们还发现了崩溃式破坏，即高一级技术代替之前的功能执行者。这意味着那些仅用于过时技术的电路不再被需要了，因此它们自己也被替代了，于是便产生了崩塌——我们可以对此加以研究和测算。[8]

我们可以以这种方式检视运动中的技术进化，它证实了我之前在本章得出的结论。

技术进化与生物进化的比较

我想对我在这章所描述的进化形式进行一些论述。首先，技术的进化根本就没有可以预先决定的确切顺序。我们无法事先就知道哪个现象会被发现并转化为新技术的基础；也无法在巨大的组合可能性中事先指出哪种组合会被创造；无法知道当这些被实现的时候，哪些机会之门会被开启。作为这些不确定性的结果，技术体的进化具有历史偶然性。它依赖于历史上的小事件：谁遇见了谁，谁借用了什么理念，哪个权威宣布了什么消息。这些小的偏差并不随时间均匀出现。它们内置于技术体中，而且由于新加入的技术依赖于已有技术，它们会将偏差进一步地传播开去。如果历史可以重演，我们或许会捕获一些相似的现象，也因此会发展大致相同的技术。但是它们出现的顺序和时机将可能不同，这将导致经济和社会历史变得与现在不同。

这并不意味着技术的进化是完全随机的。未来10年将会得到发展的技术会被理性预期，当下技术的改进也会或多或少地被预见，但是总体而言，就如同在遥远的未来，物种的集合是不能借由现在的物种所预见一样，技术体在未来经济中也是不可预见的。我们既不能预见会形成何种组合，也不能预见何种机会利基会被创造。由于潜在组合是成指数式增长的，这种不确定性也会随着技术体的发展而增加。在3 000年前，我们可以自信地说，100年后的技术将不会有太大的变化，而现在我们将很难预计50年后的技术会是什么样的。

我的另一个观点是，这种进化在时间上不是均匀分布的。有些时候，这类系统是静止的，那时机会利基是被满足的，其他元素则安静地等待着合适的组合。小的创新时有发生——时而这里有一种新的组合，时而那里有一个组合替换。其他时候，系统会处在剧烈的变化中。一种影响巨大的新组合可能出现（例如，蒸汽机），一阵一阵的变化会被释放出来。新的机会利基形成了，新的组合出现了，许多重新排布也出现了。变化导致变化，而在这期间，静止衍生静止。

在本书中，我反复提到技术体的“进化”。我是认真的。如同生物进化一样，技术的进化也很复杂。因此，在一个漫长时间的尺度上，我们可以给出另一个范例。它虽不是亘古永世，但依然是人类存在以来漫长的年代。

与生物进化相比，技术的进化到底为何？有什么不同吗？

我们的机制（组合进化）关乎新事物的创造，这种创造要通过自身组合来进行。如果要在生物进化中选择一种单纯的对应案例，则应该是选择一个被证明很有效的器官，我们假设他有一个来自鬣蜥的器官和一

个来自猕猴的器官，然后把这两个器官组合，并和其他器官一起创造出新物种。这看起来很荒唐，但是生物的确会通过时不时的组合来形成新的结构。在原始细菌中的基因通过一种叫作水平基因转移的机制来进行交换和重组。基因调节网络（大致可以理解为决定一个器官中的基因被开启的顺序的“程序”）不时地通过组合进行扩展。

我们更为熟悉的是，更大的物种结构是随着简单结构的组合而创造出来的。[9]简单结构经过组合后，真核细胞就出现了。当单细胞生物组合在一起后，多细胞生物就出现了。在生物中，功能特征（比如说，听小骨的组合形成了将耳膜连接到听神经的机制）可以从周围的零部件中形成。“在我们的宇宙当中，”分子生物学家弗朗索瓦·雅各布（Francois Jacob）说，“物质是经由陆续的整合被安排在某种结构等级当中的。”这是正确的。生物界也同样存在组合进化。

但是，这种大型组合结构的创造在生物进化中是少见的，和技术的进化相比，更少见。它不会每天都发生，而是在几百年间才会间或发生一次。这是因为生物进化一定要突破达尔文学说的瓶颈。组合（至少对于高等级的生物来说）不能从不同的系统中来选取元素并将它们一口气组合在一起。它被基因进化的限制所包围：一个新的组合的创造步骤必须是递增的；其中每个环节必须能产生可以成活的东西，即某些生物；而且每个新结构的合成都必须借助以前存在的元素。在生物中，组合会有，但不是定例，不会经常有，也不会像我们在技术中看到的那样直接。在生物进化中，变异和选择是第一位的，而组合的发生是非常偶然的，但常常会有惊人的结果。

相对来说，在技术当中，组合则是常态。每个新技术和新的解决方案都是一个组合，而且每个现象的捕捉都会应用一个组合。在技术当中，组合进化是第一位的，是常规定例。达尔文进化的变异和选择也存在，但是它们是靠后的，是在已经形成的结构之上产生作用的。

这种技术自我创造的阐述会给我们一个不同的感觉。我们开始感到技术体似乎是有血统的，它似乎是各种事件的巨大载体。这个载体能产生事件，能将事件收入其中，又能让它从中消失。这个过程既不统一也不均匀，它表现为爆炸性的创生和雪崩式的替换(12)。它持续性地探究未知，持续性地揭示新现象，持续性地创造新颖性。它是有机的：新的表层附在旧的表层上，创生和替换相互重叠。在集合的意义上，技术不仅是单个技术的总和，还是一种代谢化学，一种几乎是无限的实体的集合。它们相互作用产生新的实体以及进一步的需求。我们不该忘记，是

需求驱动着技术的进化、新组合的可能性和对现象的发掘。没有未得到满足的需求，就不会有什么新的东西出现在技术中。

最后一个想法。我说过技术是自我创生的，它是一个自我编织的活网络。那么我们是否可以从某种角度，或某种字面意义上说，技术是活的？

虽然现在没有关于“生命”的正式的定义，但是，我们在判定什么是“活的”的时候，会看它是否达到了某些标准。用系统语言（systems language），我们可以问：实体是否是自组织的，或有没有某些简单的规则使它可以自行聚集起来？实体是否是自生的，或它是自我产生的吗？看看这两条标准技术（体）是否都达到了。用通俗的语言，我们可以问：实体能够繁殖、生长、反应和适应周围环境，摄入和释放能量以保持自身的存在吗？技术（体）也通过了这些测试。它像一个组织中的细胞一样，在个体元素的意义上实现繁殖、死亡和替换。它的元素会生长，它们永不止息地生长。不论是技术体还是单个技术元素，都与它的周围环境密切结合。它和环境交换能量：它摄入能量以形成和运转每个技术，每个技术又都释放物理形式的能量回馈给环境。

以这些标准来衡量技术，它的确是生物体。但是它只是在珊瑚礁意义上的有机体。至少在技术发展的目前阶段，技术的建构和繁衍还依然需要人类作为其代理人——在这一点上我还是很欣慰的。

技术思想前沿

有生命的技术：一方面，技术是自组织的，它可以通过某些简单规则自行聚集起来；另一方面，技术是自我创生的。通过这些来衡量技术，技术确实是有生命的，不过它们只是珊瑚礁意义上的有机体。

THE NATURE OF TECHNOLOGY

第10章 技术进化所引发的经济进化

众多的技术集合在一起，创造了一种我们称之为“经济”的东西。经济从它的技术中浮现，不断从它的技术中创造自己，并且决定哪种新技术将会进入其中。每一个以新技术形式体现的解决方案，都会带来新的问题，这些问题又迫切需要进一步得到解决。经济是技术的一种表达，并随这些技术的进化而进化。

我们已经在第9章非常直观地看到了技术的进化。另一种感知这种进化的方式，就是通过经济的眼睛去观察它。一面经济的镜子会映射出技术的变化，以及我前面谈到的技术的增长和替换。经济的改变不仅是通过重新调整生产和消费模式，或者仅仅通过我们在第8章看到的创造新组合的方式来完成的，经济实际上会随着技术的进化而完成结构上的变化，即改变制度安排的方式，而这种改变可能会随时随地、在所有层次上发生。

因此，我想重新回到技术进化的每个步骤，去看看它们在经济中究竟是如何表现的。我们将会看到的是经济结构变化的自然过程，这一过程是被我们刚刚探究的那个进化过程所驱动的。首先，我们需要通过一个不同于标准经济学的方式来进行思考。

经济就是技术的一种表达

不论是在词典中还是在经济学教科书中，标准的经济定义都是将其作为产品或服务的一个“生产、分配和消费的系统”。[1]我们将经济作为某种背景，它向发生在其中的事件或某种调整提供原材料。这样看来，经济就变成了某种类似盛装技术的大集装箱，变成了一架带有许多模块和零部件的巨型机器，这些模块或零部件都是技术或生产方式。当一个新技术（例如，铁路交通）到来时，它为某个产业提供了一个新的模块，一种新的升级。它所替换的旧模块（运河）会被去掉，新的升级模块不知不觉地安插进来，机器的其他部分会自动地重新进行平衡，它的张力和变动（价格、生产和消费）也相应地进行重新调整。

这种看法并不错。我在大学中就被灌输了这样的经济学概念，而且它也是今天的经济学教科书中所描绘的经济定义。但我认为这并不完全对。为了探究结构变化，我希望能以一种不同的方式看待经济。

> 技术思想前沿
>
> 众多的技术集合在一起，创造了一种我们称之为“经济”的东西。经济从它的技术中浮现，并不断从它的技术中创造自己，并且决定哪种新的技术将会进入其中。经济是技术的一种表达，并随这些技术的进化而进化。

我将经济定义为一套安排和活动，一个社会将借助它来满足需求。（这成为经济学研究的主要内容。）关键是，这些安排是什么呢？好吧，我们将从维多利亚时期经济学家的“生产手段”（处于经济核心的工业生产过程）开始。的确，我的定义不会令卡尔·马克思感到吃惊。马克思是从“生产工具”开始审视经济的，其理论涉及了他那个时代的大型工厂和纺织机械。

但是我将从马克思的出发点（工厂和机械）再往后一些来开始我的讨论。构成经济的整套安排将包括所有制度和方法，以及所有的我们称之为技术的目的性系统，包括诊所和外科手术、市场和定价系统、贸易安排、分配系统、制度和商业，还包括金融系统、银行、监管体系以及法律系统。

上述这些都是安排，都是手段，我们通过它们实现我们的需求，达到人类的目的。因此，所有这些都是我前面给出的“技术”的定义或者目的性系统。我是在第3章谈到这点的，大家可能有点忘了，它意味着纽约股票交易所以及合同法的特殊条款都是实现人类目的的手段，这与钢铁厂和纺织机械厂没有什么两样。它们也是广义上的技术。

如果我们将这些“安排”都包含在技术的集合之中，我们就开始不再将经济看成是技术的载体，而认为它是建构在技术之上的。经济是由技术作为中介（覆盖）的一系列关于商品和劳务的活动、行为或流动，也就是说，方法、过程和组织的形式构成了经济。

经济是其技术的表达（expression）。

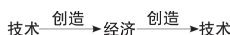
我并不是说经济和它的技术是完全相同的。对经济来说，还有许多其他类型的活动，比如商业上的策划、投资、竞标以及交易。这些都是活动而不是目的性系统。我想说的是，经济的结构是由技术铸造而成的，也可以说，技术构成了经济的框架。经济中的其余部分，如商业活动、交易中不同参与方的战略和决策，以及随之而来的物流、服务流和资金流，它们则构成了经济体的神经和血液。但是这些部分只是外围的环绕物，并且它们归根结底也是由构成了经济结构的技术（那些目的性系统）形塑而成的。

我们需要在这里做的思维方式的转变并不是很大，但很微妙。就如同不再把心灵看作容纳概念和惯性思维过程的容器，而是把它当作这个容器的产物；或者不是将生态环境看作容纳物种的集装箱，而是把它当作这些物种集合的产物，经济也是如此。经济因其技术而形成了一种生

态，因此经济形成于技术。这意味着经济不可能独立存在。如同生态一样，经济为新技术形成机会利基，当新技术产生时，这些机会利基就会被填补。

这么一想，结果就不同了。它意味着经济浮现或萌发于它的技术当中；它意味着经济不仅会随着技术的变化而重新适应，它还随着技术的变化而继续构成和重构；它还意味着经济的特征（形式和结构）也将随着它的技术的变化而变化。

总结起来，我们可以这样说：技术集合在一起，然后创造一个结构，决策、活动、物流、服务流都发生在其中，由此创造了某种我们称之为“经济”的东西。经济以这种方式从它的技术中浮现出来。它不断地从它的技术中创造自己，并且决定哪种新技术将会进入其中。注意这里的因果循环：技术创造了经济的结构，经济调节着新技术的创造（因而也是它自身的创造）。



但是，我们通常看不到这种循环，因为短期内经济是以固定的方式出现的，它呈现为一系列活动的容器。只有跨越几十年的观察，我们才能看到使经济得以形成、交互作用以及再次崩溃的安排和过程，而只有在更长的时间跨度内，我们才能看到经济的这种持续的创造和再创造。

结构性变化

当新技术进入经济时，会对这个从技术中产生的系统做些什么呢？当然我们会看到如我在第8章中谈到的调整和新组合，这些动作都会对经济造成很大的影响。除此之外，我们还会看到，新技术在推动经济结构和制度安排方面也将带来一系列的变化。

我们现在来到了一个经济学理论通常不太介入的地方，即结构性变化过程。但是这块地方并不是空的。历史学家一般会在这里驻扎，对我们来说，应该是经济史学家。历史学家看待新技术介入经济时，不仅会看到它们引起的经济的重新适应及增长，而且还会看到它们能够令经济竞争本身在它的结构中发生变化。通常，他们往往是基于专门的案例进行理论推演。相比之下，我们关于经济和技术的思考向我们提供了一个对结构变化进行抽象思考的机会。

在实践中，一项新技术很可能会创造一个新产业；为此它可能需要建立新的制度安排；可能会引发新的技术和社会问题，并因而创造出新的机会利基；而这又可能引起进一步的组合变化。如果我们借用第9章关于技术进化的步骤，我们还可以看到这种变化的顺序，只是现在我们是从经济的视角来重新看待它们而已。好的，让我们开始吧。

现在我们假设一个新技术进入了经济，它将替换旧的技术（或者过时的产业及其他依赖这些技术的技术）来实现同样的目的。当然，如我之前所说，新技术也将引起经济的某种调整。如前所述，到目前为止，这些步骤包括第1、2、4步和第6步，但是也会经历第5步和第3步：

- 新技术提供了潜在的新元素，它们可以交互地用在其他技术中。因而它担当了引导使用它、适应它的技术的角色。特别是，它可能引发包含它在内的新的组织制度的发生。
- 新技术可能为进一步的技术建立机会利基。这样的机会以不同的方式建立起来，但是它们会从新技术所引起新的技术的、经济的或者社会的问题中建立起来。因此，新技术带来了需要进一步以技术来解决这些问题的需求。

我只是用经济学术语重述了那些通过技术进化而形成机制的步骤。但是重点有些不一样，我重新转化这些步骤，然后去重点描绘的是一个产业的经济结构是怎样构成的。当一个新的技术进入经济，它会召唤新的安排——新技术和新的组织模式。新的技术或新的安排反过来会引起新的问题（或者通过现存技术的修正来适应目的），新的问题又引起了新的技术需求。所有的变动都以“问题与解决-挑战与回应”这样的序列进行，而这正是我们所谓的结构性变化。经济以这种方式在暴风骤雨般的变化当中构成或重构自身，包括创新和适应这种新的制度安排，以及如影随形、互相追逐的机会利基。

让我以一个具体的例子来详细阐明我的观点。当实用的纺织机械于18世纪60年代在英国诞生时，它提供了一种替换当时以手工作坊为基础的生产方式。[2]当时，羊毛和棉布的纺织都是在家里手工完成的，属于散工制的组织方式。新机器在刚开始时获得了部分成功。接着，它需要比家庭手工作坊更大规模的组织，因此它为更高级的组织安排（纺织厂或毛纺厂）提供了机会，并成为其中的一部分。工厂实际上也是某种技术，它作为一种组织手段，反过来需要另一种手段来实现：工厂劳动力。劳动力当然早已在经济中存在，但是当时当地并没有足够多的人来

支撑这种新的工厂系统，因此需要劳动力的调集。由于大多数劳动力来自农村，这就需要工厂附近有居住的地方，工人宿舍和工人住房建设因此而出现。随着工厂、工人以及他们的栖息地组合而来的是一个工业化城市的生长，一套新的社会组织方式或者说一套新的安排出现了，内含这种结构的维多利亚式工业经济就此开始呈现。就这样，一个时代的特征，一套和超级工业机械相匹配的安排显现出来了。

但是这个过程还远远没有结束。工厂的劳动者当中许多都是孩子，经常在狄更斯作品中那样的条件下工作，这也引发了强烈的改革需求，这种改革需求不只是关于“下层阶级的道德状况”[3]，还有他们的安全状况。继而，法律系统需要采取进一步的安排来对此作出回应，劳动法就在这种情况下被制定出来，以阻止最坏的恶行发生。然后，新的工人阶级开始要求分享工厂创造的财富，他们利用一种手段来改善他们的状况，那就是工会组织。工厂中的劳动力要比家庭作坊中的劳动力更容易组织，[4]因此，若干年之后，工会就变成了一股不可小视的政治力量。

在这种机制下，纺织机械并不仅仅是替换了手工生产方式，它还为更高一级的制度安排（工厂制度）提供了机会。在这个制度安排中，机械只是其中的一个组分。新的工厂制度反过来建立了一个服务于劳动力及其栖居需要的需求链条。反过来，对这一问题的解决又创造了进一步的需求，所有这些最终演变为维多利亚式工业体系。这个过程用了100多年的时间才真正完成。

读者可能会反对说，这样的描述可能会使结构性变化看起来过分简单了，或者说，太机械了。技术A提出了一个对安排B的需求，这个安排由技术C实现了；C又提出了进一步的需求D和E；它们又由技术F和G来完成。当然，这样的序列确实确立了结构性变化的基础，但同时，它们绝不简单。只要我们想一下这些新的安排和技术自身所引起的次一级技术和次一级安排的递归循环图景就能理解了。工厂体系需要新的动力方式应对机器，需要绳子和滑轮系统来传输动力，需要获取和追踪材料的方法、新的记账方式、管理方式，以及产品运输方式。所有这些又各自需要从其他方面建构出来，从而又产生了它们各自的需求。结构变化是不规则的碎片，它会在次一级的层级上进行进一步的分化，就像胚胎动脉系统会随着发育分化出小动脉和毛细血管一样。

技术结构性变化中的某些回应根本不是经济性的，比如手工艺的机械化会从纺织行业蔓延到其他行业。从心理学上看，工厂不仅创造了新的制度安排，而且需要新型的人力资源。历史学家大卫·兰登（David

Landes）说："工厂的运转规则需要并最终创造出了新型的劳动者……从此，纺纱人不再能在家里转动她的纺轮，织布匠也不能在家里抛动他的梭子，那种没有任何监督的美好时光从此不再。现在，工作必须在工厂中完成了，以一种无聊的、毫无生气的机器所规定的步调，作为团队的一员，要与整体共进退，开始、间歇、停止——所有这些都在密切监督之下进行，工厂用道德、经济奖惩来促使工人勤奋地工作，有时候甚至是身体上的强迫。工厂是一种新的监狱；时钟是一群新的狱卒。"[5]新技术不仅引发了经济的变化，它还引发了心理的变化。

在谈到结构变化时，我们需要知道的是，不是所有变化都是有形的或者是"被安排"出来的。我们要记住，这种变化可能是多因多果的。但同时，我想强调的是，我们可以将结构变化的过程逻辑化或理论化，摆出步骤，描述技术的进化。经济中的结构变化不仅是新技术的副产品或者仅仅是替换旧的技术，也不仅仅指紧随其后的经济的调整，它是一个后果的链条，在这个链条中，构成经济结构框架的安排，召唤新的安排。

对于制度安排的到来或者界定经济的结构来说，这里当然没有什么是必然的，没什么可以事先预料。[6]我们早就看到了有许多不同的组合、许多种安排可以用来解决技术提出的问题。许多选择取决于小的历史事件：问题出现的顺序、个性偏好或行动偏好等。换句话说，选择取决于机会。技术决定经济结构以及由此产生的大部分世界，但是哪个技术获得选择却无法事先知晓。

解决方案带来的新问题[7]

我曾经说过，结构的展现是一个连续、重复制造这种构成经济安排的过程，一套安排构成了下一套安排所需要的条件。这种复制一旦开始，就没有理由停下来。哪怕只有一项新技术（想一想计算机或者蒸汽机），一旦开始就可能绵绵无休地持续下去。

反过来，经济也从未停滞过。经济结构在大多数时间里都处于高度兼容状态，因此好像没有发生什么变化。但是，就像熊彼特在100年前指出的那样，在这种静态中，蕴藏着自我突破的种子。[8]其根本原因是创造出了新的组合、新的安排，或者如熊彼特所说的"新的产品、新的生产或运输方法、新市场、新的工业组织"[9]，形成了一个"工业突变"的过程。这种突变"从内部持续彻底地改革经济结构，不断地破旧立新"。

系统在其内部总是随时准备好进行变化的。

但是和熊彼特比起来，我这里所给出的观点要有更多的暗示。即将到来的新技术不是仅仅打破静态，比如说发现比我们现在用的产品或方法更好的新的组合：它需要一系列安置新技术的条件，并解决由此产生的另外一些新问题。如此一来，又要为新的能够解决这些问题的进一步的技术创造新的机会利基——接下来又产生进一步的问题。

因此，经济总是存在于永恒变化之中，存在于一种永恒的新颖性之中。它永远存在于自我创造的过程之中。它永远不满意。我们可以对此加以补充，在任何层面上，有意义的新技术都可以随时独立地进入经济。结果不仅是熊彼特意义上的打破均衡，而且同时还会引起连锁性变化，所有变化都重叠在一起，相互作用并引起进一步的变化。可以说，变化招致变化。

奇怪的是，我们可能对这种随时随地都会发生的同时性变化并不是很在意。这是因为结构变化的过程动辄要几十年，而不是几个月。这很像就发生在我们脚下的土地表面的缓慢隆起一样。简而言之，结构有高度的连续性，它是松散的，可兼容的系统，在这个系统中可以做计划以及采取行动。但是这个结构又随时在变化。经济永远都在建构它自身。

技术的这个不断进化的过程和经济的不断重建过程会告一段落吗？原则上，可以。但只是原则上。止步不前的情况只有在未来没有新现象被发现的情况下才可能发生，或者只有当进一步组合不知为什么耗尽了，或者当我们人类的需求用我们现有的技术就可以得到满足时才可能发生。但是这些都不太可能。永远开放的需求和总有可能被发现的新现象足以驱动技术永远向前，而经济将会如影随形。

还有一个原因令技术的进化不会停止。我一直强调，每一次以新技术作为解决方式都会创造出新的挑战、新的问题。这是一个通则：每个技术都包含着问题的种子，而且通常是很多颗。这不是技术或经济的“法则”，更不是世界的“法则”。它只是一种基于人类历史的普遍性的经验观察——一种遗憾的观察。以碳为基础的化石技术带来了全球变暖；核能（一种环保的清洁能源）带来了核废料丢弃的问题；航空交通带来了在世界范围内加速感染蔓延的可能性。在经济中，解决导致问题，问题趋向进一步的解决，这个在解决和问题之间的舞蹈在未来任何时候都很难被改变。如果幸运的话，我们会收到我们称之为“进步”的“净利润”。不论进步是否存在，这种舞蹈导致的持续变化都是由技术带来的，而经济只是一个结果。

我在这一章一直是从经济的角度来看待技术的进化。因为经济是其技术的一种表达，它是一套安排，包含着集合进化的过程、组织、装置以及制度供给。它随着技术的进化而进化。因为经济是源出于技术的，它在技术的自我创造中获得传承，永远开放、永远新颖。因此，经济最终产生于创造技术的那些“现象”之中。归根结底，那个被组织起来为我们的需求服务的东西是“自然”。

这样的经济毫无简单性可言。安排一个一个建立起来：法律体系的商业部分是在市场和合同存在的设想之下建立起来的，而市场和合同则假设银行和投资机制是存在的。经济因而不是均质的。它是一种结构，一种宏大的结构，它是互动的、包含着相互支撑的安排，存在于任何等级，历经几个世纪，不断从自身中长出自己。它几乎就是一种生物，或者至少说就是进化之物，它持续地变迁着它的结构，新的安排创造了新的可能性或问题，需要新的反应，然后是进一步的安排。

这种结构进化是对构成经济的安排的不断地重建，一套安排为下一个将要到来的安排设立条件。它与在给定的安排或行业当中进行重新调适是不一样的，而且它和经济的增长也是不一样的。它是持续的、分形的（fractal）、不可阻挡的。它会带来永不停止的变化。

那么结构变化中有什么是永恒不变的吗？经济在建构它的模式时总是汲取同样的元素：人类行为偏好、论述的现实基础、买和卖的商品相等的自明之理。这些在基本“法则”之下的元素总是相同的。但是它们所用的手段却表现为随时间而变化，它们构成和重构变化的模式也随时间而变化。每一种新的模式，每一套新的安排都会产生一种新的经济结构，就算结构消逝了，那些背后用来构成这些基本法则的元素也从来都一样。

经济学作为一门学科，经常为人诟病，因为，不像物理或者化学这些“硬科学”，经济学不能保持一套不变的描述。但是这不是经济学的失败，而是正当和自然的。经济不是简单系统，它是一个进化的复杂系统，其形成的结构永远随时间而变化。这意味着我们关于经济的解释也一定是随着时间而变化的。我有时候想，经济就像是夜色下的战场，漆黑一片，在战壕外几乎什么也看不到。而大约半英里远，就有敌人的营地，可以听到窸窸窣窣的声音，可以感到军队正在重新部署。（当然，新的战略部署即使再好，也是在现有部署的基础上转化来的。）这时，突然某个人发射了一颗照明弹，火光照亮整个战场的部署，各种炮位、安排、部队、战壕，一下子变得一目了然，然后火光骤然熄灭，一切又

复归黑暗。经济就是如此。经济学中的光芒就是亚当·斯密、李嘉图、马克思、凯恩斯，还有熊彼特的理论。它们偶尔照亮了一下战场，但是真正的骚动和变换一直在黑暗中进行着。我们确实能够观察到经济，但是我们描述经济的语言，我们标明经济的标志，以及我们对经济的理解，都凝固于那燃亮的一瞬间——尤其是最近的一组照明弹之中。

THE NATURE OF TECHNOLOGY

第11章 我们的立场是什么

随着基因组研究和纳米技术的发展，生物正在变成技术。与此同时，从技术进化的角度看，技术也正在变为生物。两者已经开始相互接近并纠缠在一起了。我们需要和自然融为一体。如果技术将我们与自然分离，它带给我们的就是死亡。如果技术加强了我们和自然的联系，那就是它对生命和人性的厚爱。

我一开始就曾说过，这本书的目的是要建立一个关于技术的理论——“一组连贯的一般命题”，并希望借此能够提供一个帮助我们理解“技术以及技术怎样在这个世界存在”的框架。我尤其希望建立一个技术自己的进化论，而不是一个从技术之外借用的理论。但是这样做的结果，用达尔文的话来讲，将会带来一场“持久的争论”，为此付出的代价将是不断地进行重述，那么现在就让我先就这个理论再做一次扼要概括吧。

理论通常始于一般命题或原理。我们的理论始于三个原理：一切技术都是元素的组合；这些元素本身也是技术；所有技术都利用现象达到某个目的。其中第三个原理特别提示我们，技术的本质就是对自然的编程，它是一种对现象的捕捉，并驾驭这些现象为人类的目的服务。某个个体技术对许多现象进行“编程”，并精心安排策划这些现象，最后使它们能够密切配合以完成特定的目的。

一旦新技术（单个技术）诞生了，它们就立刻成为可供进一步地建构更新技术的潜在构件。这个过程导致技术的发展呈现出一种进化的形态。准确来讲，是一种组合进化的形态。当然它的机制与达尔文进化论不同。首先，创造新技术的构件本身也是技术，创造出的新技术本身又成为未来更新技术的构件。其次，使得这一进化得以进行下去的是持续地捕捉和驾驭新现象，但这种捕捉和驾驭本身又需要现有技术来完成。这段叙述是在表明，技术自己创造了自己。以这样方式，技术，这种与文化相适应的机械艺术的集合体，自展式地前进着，其构成元素从仅有的几个模块发展到许许多多的模块，而模块本身也从相对简单的形式进化成极其复杂的形式。

这种进化听起来很简单，但如同达尔文进化论一样，其实不然，它包含了很多细节和机制。其核心机制当然是能够产生根本性新技术的那一个。技术的新“物种”诞生于需求和现象的链接。这种链接又表现为一个过程：首先需要假想一个概念，就是想出一组可用的效应；然后找出实现这个概念所需的元器件或集成件；最后进行组合。这个过程本身是递归性的。这是因为将一个概念放到现实中来会带来很多问题，而解决这些问题会带来更多的次生问题，因此这将是一个在不同层次的问题和解决方案之间来回跳跃直至完成特定目的的漫长过程。

这一过程的核心是组合，即将合适的组件及功能组合在一起，形成一个解决方案。这个过程既是物质的，也是精神的，但是它不是技术进

化唯一的动力。另外一个动力是需求，一种对新的做事方式的需求。而从技术本身产生的需求要多于来自人类的需求。需求动力主要来自于技术本身遭遇的极限以及存在的问题，而这些问题的解决反过来又必须通过进一步的技术来获得。因此，需求追随着解决方案，解决方案又反过来追随着需求。组合进化中营造出的需求与建构出的解决方案在数量和重要性上是不分伯仲的。

所有这一切发生的过程既不统一也不顺畅。技术集合通过采用或者丢弃某些技术，创造某些机会利基，以及揭示一些新现象来实现进化。

技术体在狭义的持续发展意义上也在进化：它们从初显到形成，时不时变换着“词汇表”，一直到最后被吸收到经济产业当中。同时单个技术也在进化——或者说发展，它们会持续变换内部的组件，或添加更复杂的集成件以达到更好的性能目标。

于是，技术在所有层次上都存在着持续的变动，在所有层次上都会间或出现新组合、添加新技术、淘汰旧技术。技术以这样的方式不断地探索未知领域，不断地创造出进一步的解决方案和进一步的需求，因之而来的是持续不断的新颖性，整个过程就此呈现了某种有机性：新技术不断在旧技术之上衍生出来，其中创造和替换交叠着推进整个过程向前发展。在集合的意义上，技术不仅是单个技术零件的总和，还是一个新陈代谢的化学过程，一组几乎无限多的实体相互发生作用，从而产生新的实体和进一步的需求。

经济导演并调和着这一切。它发出需求信号，检验商业理念的可行性，提出对新版技术的需求。但是经济不仅是技术的接收器，也不仅是升级技术的接收器，更重要的是，经济是其技术的表达。在它的结构中包含着一系列相互支持的安排——商业、生产资料、制度、组织，而这些实际上就是广义的技术本身。商业活动和行为都是围绕这些安排而发生。这些“安排”创造了“进一步安排”的机会，结果就是一个接一个的经济结构的转型。经济作为这个过程的结果，继承了技术的所有品质。如果从大的时间跨度来看，经济也是变化不羁的，如同技术一样，经济也是开放的、历史依赖的、层级式的、非决定性的，而且是不断变化的。

技术具有生物属性，反之亦然

我的观点可能会遭到反驳。我在本书中给出的许多例子都是从20世纪和19世纪的技术中挑选出来的，因此都是历史性的。那么在面向未来

技术的时候，我所认为的“技术是具备化学功能的，可以通过编程将不同结构用于满足不同目的”这种观点应该是有不对的可能性的。

但实际情况正好相反。随着技术的进步，我所描述的这种景象将变得愈加合理。数字化时代的到来实际上拓宽了组合的可能性，这是因为即使功能组件来自于不同的域，但它们一旦进入到数字化的域，就都变成了相同类型的操作对象——数据字符串，因此马上就可以用同样的方式进行组合了。远程通信技术使这些数字化元器件之间的远程组合成为可能，因此可执行程序几乎在任何地方的虚拟世界里都可以产生互动。尽管目前的传感装置还很原始，但是应用这样的系统，现在就可以完成探查周围环境并完成适当回应的任务了。上述的技术发展使得来自于不同域的、分布在不同地方的功能组件可以被临时组合在一起，形成一个暂时性网络。这种相互连接的、可以完成“物-物对话”（thingsin-conversation-with-things）的技术集合因此能够感知环境并对其进行反应。这样看来，现代客机导航系统就是一套功能组件的组合，主要包括同步罗盘系统、全球定位系统、导航卫星及其地面接收站、原子钟、自动驾驶仪、远程遥控系统、位于控制面板的制动器。这些组件之间相互对话、相互查询、相互控制、相互执行，就像计算机程序算法的一套子程序之间在相互查询、激发并执行一样。

目前，具有代表性的技术已不再是那种被固定起来的、一台机器实现一个固定的功能的模式了。现在的主流技术是一个系统或者一个功能网络，一种“物-执行-物”（things-executing-things）的新陈代谢，它能感知环境并通过调整自身来作出适当反应。

理论上，我们可以由上至下来统筹设计这个物–执行–物的网络，但实际上这非常困难。当一个网络是由数以千计各自独立又相互作用的元器件组成时，加上迅速变化的环境，这时几乎不可能有任何可靠的方法来进行这种自上而下的整体性设计。因此，越来越多的网络被设计成可以从经验中进行“学习”的网络，主要是去学习不同环境下各种元素之间如何作出互动的简单规则。学会这些规则，它们就可以对感应到的环境作出恰当的反应。那么，这种设计的结果构成所谓的“智能”了吗？某种程度上，是的。一个简单生物的认知，比如说E型大肠杆菌“觉察”到葡萄糖浓度的增加并向那个方向移动，可以被定义为“能够感知环境并适当反应”。因此，当现代技术逐渐进入到一个网络中，能够感知、配置、恰当地执行，它就表现出了某种程度的认知能力。从这个意义上说，我们正在向智能系统前进。基因技术和纳米技术的到来将加速这一进程。事

实上，未来这样的系统不仅能够自构成、自优化、具有认知能力，还能自集成、自修复以及自保护。

我在这里并不想谈论某种科幻未来，或者讨论这种趋势将意味着什么，因为会有人去做这件事的。我想请读者注意的是，在过去，诸如自构成、自修复以及认知能力这样的词语会用在什么样的事物上面呢？过去我们是不会把它们用在技术上的，它们都是生物学词汇。现在，这些词汇被用来描述技术，是否能够表明随着技术复杂程度的提高，技术将具有生物特征呢？这看起来很矛盾，因为技术的本质当然应该是机械性，因此当它变得更复杂时，它应该只是具有更复杂的机械性而已。那么，技术是怎样变得具有生物特征的呢？

答案有两个。一个是所有的技术在某种意义上都同时具有机械性和生物性。如果你从上至下地检视一项技术，你可以将其看作是一种互相连接的元器件的安排，这些元器件为达到某个目的而互相制约、互相啮合。因此技术变成了像钟表一样的发条装置，它变得具有机械性了。但是，如果你在思想中将它从下往上进行检视的话，想象这些零部件是如何聚集到一起的，你则会将它们看成一个整体，一个完整的器官，一个更高等的、具有功能性、目的性的整体。它变成了一个机体。因此，一项技术具有机械性还是生物性取决于你的立场。另一个答案则强调，技术完全是生物性的。技术具备能使我们联想到生物的某些属性：当它们感知环境并产生反应，当它们变得可以自组装、自构成、自修复并且能够“认知”的时候，它们就越来越像生物了。技术越复杂、越“高技术”，就越具有生物性。我们正在慢慢开始接受，技术是机械性的，同时也可以新陈代谢。

与之相对应的一面是，当技术的生物属性变得比较好理解时，我们还是坚持认为它更具机械性。当然那种认为生物体包含着像机器一样相互作用的连接部件的观点并不新鲜，至少回溯到17世纪20年代，梅森和笛卡尔时期的哲学家们就已经开始思考“机器是否有可能和生物有机体一样”的问题了。与那时不同的是，我们现在已经了解了更多生物机体的工作机理。20世纪50年代以来，我们一点一点地梳理出DNA的精细的工作原理以及细胞内蛋白质的制造原理，部分地了解了如何精确控制基因的表达，以及大脑的部分功能。虽然这些工作还远未完成，但是它们揭示出，生物体和细胞都是非常精致的技术。事实上，有生命的东西为我们指明了科技要走的路还有多远。没有什么工程技术能像细胞的工作那样复杂。

从概念上看，生物学正在变成技术。从实际上看，则技术正在成为生物学。两者已经开始相互接近，而且，事实上，随着我们陷入更深的基因组研究、纳米技术以及其他许多技术，确实已经开始互相纠缠在一起了。

繁衍性经济

技术正在变得更具结构性和生物性，经济是否以某种方式反映了这个现象呢？如果经济是其技术的表达方式的话，那么，它一定会对此有所反映。事实的确如此。

技术从维多利亚时代那种以大宗材料加工为主导的技术模式中脱离出来，已经是很久远的事了，而现在它又开始转换了：从只实现单一目标功能的流程或机器转换为采用不同组合以实现不同目标的技术。为了反映技术的这种转换，经济至少在高科技部分更多地表现为更关注如何聚集、拼凑事物，而不是如何对现有操作进行完善。当然，商业的运作部分（例如，商业银行、石油公司、保险公司）还依然体现着大工业时代的技术特征，但是有理由相信，随着初创公司、风险投资、金融衍生品、数字化或组合生物学等越来越多，经济也必将进行以实现短期的可重构为目标的实践和商业过程的功能组合。

简而言之，经济正在变得具有繁衍能力。它关注的焦点从优化固定操作转变为创造新组合以及新的可配置的产品了。

当然，对于创造这些新组合的初创小公司的企业家们来说，这一切都不甚清晰。他们经常不知道谁会是自己的竞争对手，他们也不知道这种新工艺能否奏效、如何才能被接受，他们不知道政府会对此加以什么样的限制。一切都好像是在赌场下注，而游戏规则和收益却还不太明了。新组合周围的环境不仅不明朗，某些方面甚至根本就是未知的。

这意味着高科技经济的“问题”还没有被清晰地界定出来。如此一来，他们就无法拥有最优的“解决方案”。在这种情况下，管理的挑战不在于如何合理地解决问题，而在于要对一个未加定义的情况“找到感觉”，即去“识别”它或者把它纳入一个可以应付的框架之内，之后再对其供应部件进行相应的定位。这再次出现了悖论，技术越高级，对其的商业性处理越缺乏理性。先进技术领域内的企业家不仅仅需要决策，更重要的是，他需要在不断出错的情况下给出某种认知次序。技术思想家约翰·西利·布朗（John Seely Brown）告诉我们：“管理已经从制造产品转

移到使产品有意义上来了。”

在繁衍性经济中，竞争优势不仅只是来自于资源储备及将这些转变为最终产品的能力，而且来自将深层知识储备转化到新的战略性组合的能力。具体表现为，从占有资源的角度获取国民财富的总量开始不如从占有专业科学及技术知识的角度获得的那样多了。公司的竞争优势绝大部分是从它们占有的技术专长而来的。通常公司会缺乏可供拼凑成新组合的技术专长，竞争压力又使得他们没有足够的时间进行内部研发，它们因此常常需要购买小公司，或者与其他占有必备资源的公司结成战略联盟。结盟常常是出于某个特定的目的，然后再重新对其进行配置或者将其淘汰。这就是为什么我们会看到公司层次的组合过程常常呈现为松散联盟的持续重组——短暂，但偶尔会非常成功。

现代技术的本质发生了一系列新的转变：管理上从优化生产过程到创造新组合，即新产品或新功能；从理性到意会；从以商品为基础的公司到以技术为基础的公司；从购买要素到形成联盟；从稳态操作到不断适应。所有这些变化都不是突然发生的，事实上，新旧风格的诸要素在经济中常常是共存的。

这两个世界相互重叠，且高度相关。但是当一个“更技术性”的经济走上舞台的时候，我们就从20世纪由工厂和投入-产出关系构成的机器态经济（machine-like economy）转换到了21世纪有机的、相互联系的经济形态。如果说旧经济是一部机器，那么新经济就是一门化学。它不断创造自己，产生新的组合，总在发现着，永远处在过程之中。

经济学自身正在逐渐开始回应这些变化，并且它的研究对象不再是所谓的均衡体系，而是一个进化的复杂系统。[1]它的元素（消费者、投资者、公司、主管当局）都对这些因素创造的模式有所回应。经济学的标准理论曾经是建立在可预测性、秩序、均衡以及行为理性等基本原理之上的，其他与之相适应的应该是长时期保持不变的大批量生产方式。但是随着经济发展更趋于组合性，技术也更加开放，新的原理就会进入，并成为经济学的基本原理。秩序、封闭、均衡作为组织解释的方式现在让位于开放性、不确定性以及持续不断的新颖性的涌现。

纯粹秩序与混质活力

除了对经济变化的理解反映出一个更加开放、有机的观点，我们对世界的阐释也变得更开放、更有机了。参与其中的技术也产生了这种转

变。从笛卡尔时代起，我们开始用技术的知觉品质这类术语来解释世界：它的机械性连接、合法性秩序、驱动力量、简单的几何学、清楚的界面，它美丽得像钟表一样的精准性。这些品质投射到文化中，被认为是解释和效法的典范，并被伽利略和牛顿科学的线性次序和钟表一样的精准性特征大大加强了。我们慢慢形成了这样一种世界观：世界是由部分合理地构成的，这些组成部分由理性（Reason，在18世纪时要大写，词性为阴性）和简单性所统辖。借用建筑师罗伯特·文丘里（Robert Venturi）的说法，这样的世界观催生了一种追求纯粹秩序的古板梦想（prim dreams）。

牛顿以后的300年时间，是对技术、机器和纯秩序的漫长迷恋期。机械观逐渐占据主导地位，直到20世纪依然可见。在许多学术领域，如心理学和经济学，机械论的解释抑制了那些关于技术魅力的深刻洞见。机械观给哲学带来的希望是，它为理性哲学提供了奠基性的逻辑元素以及建构它们的语言；它为政治带来的理想是，控制一个社会并对其进行工程化操作，从而使可控制的社会结构是有可能的；在建筑领域，它带来了那种毫无装饰性的几何结构，柯布西耶（Le Corbusier）的平面（dean surfaces）以及包豪斯风格。但最后，在所有这些追求纯秩序的机械梦想破灭之后，这些领域都超出了自身的系统并无限蔓延开去了。

取而代之的观点是：这个世界反映的绝对不只是它的机械性。机械性依然是中心议题，但是我们现在认识到，随着机械变得具有互动性并复杂起来，它们所揭示的世界也成了一个复合体。它们是开放的、进化的，并且表现出了无法从某个部分预测未来的涌现性质（emergent properties）。我们所倾向的观点不再是一种纯粹的秩序，它是一个整体，一个有机的整体，而且是不完美的。[2]又是文丘里，他在谈到建筑时写道：

> 我喜欢的元素是杂交的而不是纯种的；是妥协折中的而不是一以贯之的；是曲折蜿蜒的而不是直截了当的；是模糊歧义的，而不是清晰缜密的。它们既客观又倔强，它们既无聊又有趣。它们是依惯例传统的而不是设计出来的；是随和迁就的而不是特立独行的；是冗余累赘的而不是简洁单纯的。它们既残缺不全又富于创新，是前后矛盾、模棱两可的而不是直接和清楚的。我赞同凌乱的活力优于明显的统一，我容纳不合理的结论。我赞成丰富和含义深长胜于含义清楚，我既赞同隐含的功能，又赞同外显的功能。[3]

文丘里所说的是凌乱的生命力和丰富的含义——是的，我为此倾心。我们正在将完美替换成整体，在整体之中是一片混乱的活力。这种思维的转变更多的是受进化生物学的兴起和简单机械观的枯竭的影响，而不是现代技术。但是这种影响被现代科技的特征强化了，这些特征包括连通性、适应性、进化趋势、有机性，以及它的凌乱的生命力。

我们应该怎样看待技术

我们刚才追问的是，我们应该如何透过技术看待这个世界，但是回头看，技术自身是怎样的呢？我们怎样才能看到它呢？我们应该如何看待我们的这个创造物呢？

面对正在稳步增长的技术，我们当然会深感矛盾。但这种矛盾心理并不直接地来自于我们和技术的关系，而是来自我们与自然的关系。读者完全不必为此感到惊讶，因为如果技术是为了达成我们的目标而被组织起来的自然的话，那么在很大程度上，我们与我们这种对自然的应用之间的关系将决定我们应该如何看待技术。

1955年，海德格尔做了一场演讲，题为《技术的追问》。他说，技术的本质绝不是技术的。它是一种看见自然的方式，是让所有本质上的东西自我揭示，成为人类可以加以利用的潜在资源。这是令人遗憾的。“大自然成了巨大的加油站”[4]，而且我们认为这是可利用的资源，是仅供实现我们目的所用的“储备资源”。技术或技艺，在希腊意义上的手艺或行动认知，那时被定义为“将真带入美”，那情景就如同一个古代银匠正在亲手制作一个祭献用的圣杯。技术追求的是要世界去配合它，而不是要它去配合世界。海德格尔并没有说问题存在于技术，而说它存在于随技术而来的态度之中。我们曾对大自然满怀尊重——事实上是敬畏，到现在我们已经开始“攻击”自然并将其分解成不同的部分随时供我们使用。

更糟的是，我们所创造的技术除了响应人类的需要之外，它们还有自己的需要。“技术根本不是人造的或人可以控制的工具，”[5]海德格尔的译者阐释他的话，“而是那样一种现象，它的统治来自于技术自身的存在（being），而这集中体现在西方的全部历史当中。”其他人，特别是法国的社会学家雅克·埃吕尔（Jacques　Ellul）的说法几乎与此一样，只是语言极富戏剧性。他认为，技术是一种物（thing），它指导人类的生活，[6]人类生活必须向它屈服并进行适应。技术是一种“向死而生并能自

我决断（self-determination）的有机体”，它本身即是目的。

然而，海德格尔也承认，我们和技术的关系还不错。技术创造了我们的经济以及因之而来的财富和安全。它使我们活得比我们的祖先更长久，使我们摆脱了许多他们曾面临的悲惨境遇。

认为技术控制我们的生活，或认为技术服务我们的生活，这两种观点都对。但它们也同时引起了不安，引起了一种持续的紧张情绪，这些在我们对待技术的态度上以及围绕它的政治活动中都有所表现。

这种紧张不仅因为技术迫使我们去开发自然，还来自于它决定了我们大部分的生活。之所以会这样，我在第1章曾说过，是因为对所有的人类存在来说，自然是我们的家——我们信任的是自然，而不是技术。但同时，我们仍然指望科技能照顾我们的未来——我们寄希望于技术。这样一来，我们实际上是把希望寄托在了一些我们不太信任的东西上了。这有点儿讽刺。我说过，技术是对自然的编程，是对自然现象的合奏和应用，所以在最深的本质上，它应该是自然的，是极度的自然，但它并不使人感到自然。

如果我们仅仅使用自然现象的原始形态去驱动水车或推动帆船，我们对技术就会有家的感觉，我们的信任和希望就不会那么不一致。但是现在，随着即将到来的基因工程、机器智能、仿生学、气候工程学，我们正在开始使用技术（利用自然）直接干预自然了。对于我们这种灵长类动物，对于我们这种以树、草和其他动物构成的生境为家的动物来说，这种感觉极度不自然。这扰乱了我们内心深处的信赖。

这种内心深处的不安会不知不觉地反映在很多方面。我们开始转向传统，转向环境保护主义，开始回头倾听家庭价值观，我们转向原教旨主义，我们抗议。我们这些反应的背后的实质，不论合理与否，是恐惧。我们害怕技术将我们与自然分离，我们害怕技术破坏了自然，破坏了我们的自然。我们害怕这种技术现象不在我们掌控之下。我们害怕自己曾经的什么虚无缥缈的行为在某种程度上获得了它的生命，然后它会在某种程度上反过来控制我们。我们害怕技术作为一种有生命的东西将会给我们带来死亡。不是“不存在”这个意义上的死亡，而是更糟糕的死亡，一种丧失自由的死亡，一种意志的死亡。

我们察觉到了这些，我们时代流行的神话也向我们指出了这一点。无论是在小说里还是在电影里。假如我们研究一下这些故事，我们会看到问题不在于我们是否应该拥有技术，而是在于我们应该接受冷酷的、

无意志的技术还是应该接受有机的、具有生命力的技术。在电影《星球大战》中，代表技术“恶”的是死星。那是一个巨大的东西，通过慢慢地使克隆人不断减少人性，让所有东西都置于机器的控制之下。然后，一切都消褪了颜色、去除了意志。影片的主角达斯·维德也不是一个完整的人。他的构成包括部分技术、部分人体。相比之下，英雄们，如天行者卢克和汉·索罗，则是完整的人。他们有个性、有毅力，他们在一家名叫摩斯·埃斯里的小酒吧和生物们聚会。这些生物奇怪、畸形、反常，但是充满着凌乱的生命力。再看看英雄们，他们也拥有技术，但他们的技术有所不同，这些技术不是神秘的、没有人性的。他们的星际飞船是虚弱的、有机的，而且必须踢它才能让它运转。这是至关重要的，他们的技术是人性的。它是他们自然的拓展，它易错、独特，因而也是仁慈的。他们没有用人性和技术做交易，也没有使意志向技术投降。技术向他们投降了。并且由于这样做了，技术也拓展了它们的自然性。

因此，在神话中，我们对技术下意识的反应是，我们并不排斥技术。没有技术就没有人类；技术对我们成为人起了非常大的作用。罗伯特·皮尔西格（Robert Pirsig）说：“佛陀与上帝居住在数字计算机的电路里或周期传动的齿轮中与居住在山巅或莲心中同样舒服。”[7]技术是更深的法则的一部分。但在我们的无意识中，已经把技术奴役我们的本性和技术拓展我们的本性之间进行了区别。这是一个正确的区别。我们不应该接受技术使我们失去活力，我们也不应该总把可能和想要划等号。我们是人类，我们需要的不只是经济上的舒适。我们需要挑战，我们需要意义，我们需要目的，我们需要和自然融为一体。如果技术将我们与自然分离，它就带给了我们某种类型的死亡。但是如果技术加强了我们和自然的联系，它就肯定了生活，因而也就肯定了我们的人性。

注释

前言　技术的追问

[1] 我在1987年的时候和斯图尔特·考夫曼（Stuart　Kauffman）谈论过这个想法。随后考夫曼在他的许多文章中就技术的自我创造方面作了进一步的研究。

[2] 特别参考Aitken, Constant, Hughes, Landes, Rhodes以及Tomayko的观点。

第1章　问题

[1] Meel Velliste, et al. "Cortical Control of a Prosthetic Arm for Self-feeding," *Nature*, 453, 1098-101, June 19, 2008.

[2] *Tableau Élémentaire de l'Histoire Naturelle des Animaux*, Baudouin, Paris, 1798.

[3] Stephen Jay Gould, "Three Facets of Evolution," *Science, Mind, and Cosmos*, J. Brockman and K. Matson, eds., Phoenix, London, 1995.

[4] Gilfillan, 1935a.

[5] George　Basalla在1988年的著作《技术的演化》是迄今为止最完整的理论，但最终不得不承认（p.210）"我们还无力阐明新的人工物是如何产生的"。

[6] 一个早期的例子参见Robert Thurston, *A History of the Growth of the Steam Engine*, Kegan Paul, Trench, & Co, London, 1883, p.3.

[7] Schumpeter，1912.熊彼特在写《经济发展理论》之前不久曾去瑞士访问了均衡经济学的元老Léon Walras，他告诉熊彼特："当然经济生活实质上是被动的，它仅仅是为了适应那些对其会发生作用的自然和社会影响。"参见Richard　Swedberg, *Schumpeter: A Biography*, University Press, Princeton, 1991, p.32.

[8] Usher, p.11；另见Gilfillan, 1935b, p6; McGee.

[9] Ogburn, p.104.

[10] 这个"理论"的定义源自*Dictionary.com Unabridged (v.1.1)*. Random House, Inc., accessed 2008.

第2章　组合与结构

[1] 依据《美国大学辞典》，技术是"与工业艺术有关的知识的分支"；或者按照《韦伯斯特辞典》，技术是"以实践为目的的对知识进行应用的科学"；或者按照《大英百科全书》，技术是"关于制造和做事的技艺的系统性研究"；"手段的总和"则出自*Webster's Third New International Dictionary*, Merriam-Webster, 1986.

[2] 我不同意技术即知识。知识对技术来说是必要的，比如关于如何建构、考虑、处理技术的知识，但是这些并不能使技术等同于知识。我们也可以说技术对数学中关于定理、结构与方法的知识是必要的，但是那并不能使数学等同于知识。知识是对信息、事实、理解的占有，知识对这些东西的占有不等于说知识就是这些东西。对我来说，技术是能够执行的东

西。如果你打算从飞机上跳下来，你需要的是降落伞，而不是如何制造降落伞的知识。

[3] 大多数现代半导体也包括处理外差作用的阶段，将高频率信号转换为固定中频信号以使后续电路可以优化使用。

[4] 尼采评论道："所有概念都源于我们对不等同事物的等化。没有完全相同的两片叶子，概念的'叶子'是通过对个体差异的武断抽象，通过对差别的遗忘而形成的；随之产生了一个想法：在自然中，除了'叶子们'，还应该有'叶子'——某种原型，以供编织、标识、复制、染色、弯曲、绘制形成后来所有的叶子，只是这个过程借由的是笨拙的手，因而没有一个复制品能够达到准确、可靠并忠实于原型。"参见"On Truth and Lie in an Extra-Moral Sense," *The Portable Nietzsche*, Penguin, New York, 1976, p.46.

[5] 这个想法要回溯到20世纪50年代；K. S Lashley, "The problem of serial order in behavior," in L. A. Jeffres's, ed., *Cerebral Mechanisms in Behavior*, Wiley, New York, 1951; also F. Gobet, et al., "Chunking mechanisms in human learning," *Trends in Cognitive Sciences*, 5, 6:236-243, 2001.

[6] Smith, *The Wealth of Nations*, 1776, Chapter 1.

[7] 鲍德温（Baldwin）和克拉克（Clark）阐明模块化是随时间增加的。

[8] 赫伯特·西蒙讨论过层级式系统，但是没有提到递归性。

[9] 相关的性质，即组件实体和更高层次实体相似，被称为"自相似"。当我说结构是分形的，在非严格意义上，我指的就是这个意思。严格来讲，一个分形是一个几何对象。

第3章　现象

[1] 关于物理现象的资料，参见Joachim Schubert, *Dictionary of Effects and Phenomena in Physics*, Wiley New York, 1987.

[2] Interview with Geoffrey Marcy, February 20, 2008.

[3] James Hamilton, *Faraday: The Life*, HarperCollins, London, 2002.

[4] John Truxal, "Learning to Think Like an Engineer: Why, What, and How?" *Change 3*, 2:10-19, 1986.

[5] Robert P. Crease, *The Prism and the Pendulum*, Random House, New York, 2003.

[6] "被迫在两个错误的观念中作出选择，一个是被视为异端的'科学是应用的技术'，另一个则是传统观点'技术是应用科学'。"技术哲学家Robert McGinn（P.27）

[7] Joel Mokyr.

第4章　域

[1] 当然，域通常遭遇的是集合性名词的问题。（到底谁是保守派？到底是什么构成了文艺复兴时期的建筑？）域也可能重叠。滚子轴承就同时属于几个常用的域。

[2] Tomayko.

[3] Joel Shurkin, *Engines of the Mind*, Norton, New York, 1996, p.42; Doron Swade, *The Difference Engine*, Penguin Books, New York. 2002, p.10.

[4] 出版商Dover's在1962出版的《从地球到月球》。

[5] Henry James, "The Art of Fiction," *Longman's Magazine* 4, September1884.

[6] "Preface to a Grammar of Biology," *Science* 172, May 14, 1971.

[7] 有时语法的存在并没有明显的自然后盾。像C++这样的程序语言的语法是人造的，基于一套事先商定的原则。……语法不必非得来自我们关于自然如何作用的“正式的”理解，不必非得从科学当中衍生出来。虽然大多数新语法确实是这样产生的，比如纳米技术，或者光数据传输。但是，统治一些旧技术的原理是产生于实践，产生于对自然的随意的观察，比如，金属熔炼或者皮革鞣制。

[8] 飞机规则参见Vincenti, p. 218.

[9] James Newcomb, "The Future of Energy Efficiency Services in a Competitive Environment," Strategic Issues Paper, E Source 1994, p.17.

[10] David Gelernter, *Machine Beauty*: *Elegance and the Heart of Technology*, Basic Books, New York, 1998.

[11] 引用和评论源自Annie Dillard, *The Writing Life*, Harper & Row, New York, 1989.

[12] Paul Goldberger, "Digital Dreams," *The New Yorker*, March 12, 2001.

第5章　工程和对应的解决方案

[1] 托马斯·库恩称常规科学为“正常的科学”，爱德华·康斯坦（Edward Constant）也提出了“正常的工程”的说法。我不喜欢“正常”这个术语，因为这暗示着工程也有与科学中相平行的活动，而实际上工程中并没有。我更倾向称之为标准工程。

[2] 关于波音747的引用来自Peter Gilchrist, *Boeing 747*, 3rd Ed, lan Alien Publishing, Shepperton, UK, 1999; and Guy Norris and Mark Wagner, *Boeing 747: Design and Evolution since 1969*, MBI Publishing Co., Osceola, WI, 1997; personal communication, Joseph Sutter, Boeing, November 2008.

[3] Ferguson, p. 37.

[4] 弗格森（Ferguson）告诉我们这是视觉可见的。我不是不同意，但是我说的是发生在无意识中的，不必要一定达到视觉层次。

[5] 更多参见Gelernter.

[6] Billington.

[7] Rosenberg, p.62.

[8] Neil Sclater and Nicholas P. Chironis. *Mechanisms and Mechanical Devices Sourcebook*. 4th Ed. McGraw-Hill New York, 2007.

[9] Dawkins, The *Selfish Gene*, Oxford University Press, New York, 1976.

[10] 与理查德·罗兹（Richard Rhodes）的私下交流，以及Theodore Rockwell, *The Rickover Effect*, Naval Institute Press, Annapolis, MD, 1992.

[11] Robin Cowan, "Nuclear Power Reactors: A Study in Technological Lock-in," *Journal of Economic* History 50, 541-556, 1990; Mark Hertsgaard, *Nuclear Inc: The Men and Money Behind Nuclear Energy*, Pantheon Books, New York, 1983.

[12] Malcolm Chase p.16, *Early Trade Unionism*, Ashgate, Aldershot, UK, 2000.Inner quote from L.F. Salzmann, *English Industries in the Middle Ages*, p. 342-343, Oxford University Press, 1923.

第6章　技术的起源

[1] 本章中的大部分材料来自我的论文"The Structure of Invention", *Research Policy 36*, 2-274-287, March 2007.

[2] Schumpeter, 1912, p. 64

[3] 引自Constant　p.196。除了惠特尔和冯·奥海因，其他人当然经历了早期版本的喷气发动机。

[4] Constant；Whittle

[5] Russell Burns, "The Background to the Development of the Cavity Magnetron," in Burns, ed., *Radar Development to 1945*, Peter Peregrenus, London, 1988; E. B. Callick, *Meters to Microwaves: British Development of Active Components for Radar Systems 1937 to 1944*, Peter Peregrinus, London, 1990; and Buderi.

[6] 引用劳伦斯的话源自他的诺贝尔演讲，"The Evolution of the Cyclotron," December11, 1951. Wideröe's paper is"Űber ein neues Prinzip zur Herstellung hoher Spannuneen", *Archiv fűr Elektrotechnik*, XXI, 386-405, 1928.

[7] 偶尔原理也可以通过系统性的可能性调查而获得。"我因而开始系统地检视所有可能的替换方式，"弗朗西斯· W.阿斯顿在谈到他发明的质谱仪时说道。Aston, "Mass Spectra and Isotopes," Nobel Lecture, December12, 1922.

[8] 引自Constant, p.183.

[9] Gary Starkweather "High-speed Laser Printing Systems," in M. Ross and F. Aronowitz ed's *Laser Applications* (Vol. 4), Academic Press, New York, 1980; and Starkweather Laser Printer Retrospective," *in 50th Annual Conference: A Celebration of All Imaging*, IS&T, Cambridge, MA, 1997.

[10] Townes, P.66.

[11] 关于盘尼西林，参见Ronald Hare, The *Birth of Penicillin*, Alien and Unwin, London, 1970; Trevor I. Williams, *Howard Florey: Penicillin and After*, Oxford, London, 1984; Eric Lax, *The Mold in Dr. Florey's Coat*, Henry Holt, New York, 2005; Ronald dark. *The Life of Ernst Chain: Penicillin and Beyond*, St. Martin's Press, New York, 1985; and Ernst Chain "Thirty Years of Penicillin Therapy," Proc. *Royal Soc. London*, B, 179 293-319, 1971.

[12] 关于创造性洞见背后的无意识过程的文献量有一个逐渐增长的过程，比如Jonathan Schooler and Joseph Melcher, "The Ineffability of Insight," in Steven M.Smith, et al, eds., *The Creative Cognition Approach*, MIT Press, Cambridge, MA, 1995.

[13] Townes; M. Berlotti, *Masers and Lasers: an Historical Approach*, Hilger, Bristol, 1983; and Buderi.

[14] Mullis, *Dancing Naked in the Mind Field*. Vintage, New York, 1999.

[15] 技术作家们称其为组合或积聚的观点。康斯坦应用这种观点出色地展示了汽轮机和增压空气压缩机怎样与燃气轮机的经验一起带来了涡轮喷气飞机的诞生。

[16] Charles Sűsskind, "Radar as a Study in Simultaneous Invention," in Blumtritt, Petzold, and Aspray, eds. *Tracking the History of Radar*, IEEE, Piscataway, NJ, 1994, pp. 237-45; Sűsskind, "Who Invented Radar?," in Burns, 1988, pp. 506-12; and Manfred Thumm, " Historical German Contributions to Physics and Applications of Electromagnetic Oscillations and Waves," *Proc.Int. Conf. on Progress in Nonlinear Science, Nizhny Novgorod,*

Russia, Vol.II;*Frontiers of Nonlin. Physics*, 623-643, 2001.

[17] M. R. Williams, “A Preview of Things to Come: Some Remarks on the First Generation of Computers," in *The First Computers-history and Architecture*, Raul Rojas and Ulf Hashagen, eds., MIT Press, Cambridge, MA, 2000.

[18] 大卫·雷恩（David Lane）和罗伯特·麦克斯菲尔德（Robert Maxfield）在谈到生成关系时说："它能引起的变化可以使参与者看到他们的那个世界并活动于其间，甚至会产生新的实体，比如代理、人工物，甚至制度。Lane and Maxfield, "Foresight, Complexity, and Strategy," *The Economy as an Evolving Complex System*, W . B. Arthur, S. Durlauf, and D. A. Lane, eds., Addison-Wesley, Reading, MA, 1997. Aitken (1985, p. 547)提出："去理解这个（创新）过程，必须了解先前的不同的信息流和知识储备，然后将它们聚集在一起来产生新的东西。"

[19] Simon Singh, *Format's Last Theorem*, Fourth Estate, London, 1997, p. 304.

第7章　结构深化

[1] 经济学家们称之为技术轨道（technological trajectory）。参见Richard Nelson and Sidney Winter, "In Search of a Useful Theory of Innovation," *Research Policy* 6, 36-76, 1977; G. Dosi, "Technological Paradigms and Technological Trajectories," *Research Policy* 11, 146-62, 1982; and Dosi, *Innovation, Organization, and Economic Dynamics*, Edward Elgar, Aldershot, UK, 2000, p. 53.经济学家在这里有很多话要说，而我并未谈及。他们探讨围绕着科学技术的知识库是如何影响发展的路径的，比如，通过公司的激励措施、改进不同技术体的搜索过程、专利制度和法律环境、学习效果，以及适合的工业技术结构。

[2] 用生物学的语言，我们会说这项技术在"辐射"。

[3] 关于达尔文的思路参见Stanley Metcalfe，*Evolutionary Economics and Creative Destruction*, Routledge, London 1998: Saviotti and Metcalfe;also Joel Mokyr, "Punctuated Equilibria and Technological Progress, "*American Economic Assoc. Pavers and. Proceedims* 80, 2, 350-54, May 1990; Basalla.

[4] Robert Ayres, "Barriers and Breakthroughs: an 'Expanding Frontiers' Model of the Technology-Industry Life Cycle," *Technovation* 7, 87-115, 1988.

[5] Constant, "log-jams and forced inventions" (p.245), and of "anomaly-induced"technical change (pp. 5, 244).

[6] 早期关于这一点的讨论，参见Arthur, "On the Evolution of Complexity, "in *Complexity*, G. Cowan, D. Pines, D. Melzer, eds., Addison-Wesley, Reading, MA, 1994; also Arthur, "Why do Things Become More Complex?," *Scientific American*, May 1993.

[7] Prankel, "Obsolescence and Technological Change in a Maturing Economy," *American Economic Review* 45, 3, 296-319, 1955.

[8] Vaughan, *Uncoupling*, Oxford University Press, New York, 1986, p. 71.

[9] 生物的功能变异现象类似于此。这是为了一个新的目的而使用现有的部分。比如，将手臂变异成网状以利于飞行。自适应伸展略有不同，因为更经常涉及一些系统深化，而不是将已有的部分用于不同的目的。

[10] Samuel D. Heron, *History of the Aircraft Piston Engine*, Ethyl Corp., Detroit, 1961; Herschel Smith, Aircraft Piston Engines Sunflower University Press, Manhattan, Kansas, 1986.

[11] 其他人也将技术轨道与库恩的科学发展的观点进行了对比（e.g., Dosi）。

第8章　颠覆性改变与重新域定

[1] 偶有例外。程序语言（无疑属于域）是由个人或公司将其谨慎地组织在一起的。

[2] 我必须指出，基因工程比我这里所谈论的要宽泛得多。它有许多农业上的应用，以及诸如单克隆抗体生产技术。关于它的早期发展的一个较好阐述来自：Horace F. Judson, “A History of the Science and Technology Behind Gene Mapping and Sequencing.” in *The Code of Codes*, Daniel J. Kevles and Leroy Hood, eds., Harvard University Press, Cambridge, MA, 1992.

[3] 参见Perez。本章节关于技术构建的说明多是我自己的观点。

[4] David Edgerton, *The Shock of the Old: Technology and Global History Since 1900*, Oxford University Press, New York, 2006.

[5] 罗森伯格指出，一项技术的最初用途很少是它最后的功用。

[6] 这也发生在经济增长过程中。其发生机制是一个被称为内生性增长理论的经济分支的主题。一项新技术的产生意味着达到经济目标所使用的资源要比以前少。新技术中的内生知识也能够溢出到其他产业中去。由于这两点，经济才能增长。

[7] 关于它的经济影响，参见*Robert W. Fogel, Railroads and American Economic Growth*, Johns Hopkins Press, Baltimore, 1964; Alfred Chandler, *The Railroads*, Harcourt, Brace & World, New York, 1965; and Albert Fishlow, *American Railroads and the Transformation of the Antebellum Economy*, Harvard University Press, Cambridge, MA, 1965.

[8] 休斯（Hughes）稍有不同的表述，参见“technological momentum,” in *Does Technology Drive History*, M. Smith and L. Marx, eds., MIT Press, Cambridge, MA, 1994.

[9] David, "The Dynamo and the Computer," AEA Papers & Proc. 80, 2, May 1990.

[10] 迈克尔·波兰尼（Michael Polanyi）很久以前指出，许多人类知识都是难以用语言来表达的，这样的知识确实是不可或缺的。“驾车的技能不可被关于车的理论训练所取代。”Polanyi, *The Tacit Dimension*, Anchor Books, New York, 1966, p. 20.

[11] “开放性创新”则使得专家们通过网络就可以对创新进行贡献，而不必真的在物理世界聚集在一起，这似乎与我们所说的相悖。的确，这很有用，也很重要，但它并不能提供面对面的交流，也不能轻易地形成某种文化，亦无法组织本地资源（只要走到走廊上就能利用这些资源）。可参见John Seely Brown and Paul Duguid, *The Social Life of Information*, Harvard Business Press, Cambridge, MA, 2000.

[12] Brian Cathcart, *The Fly in the Cathedral: How a Group of Cambridge Scientists Won the International Race to Split the Atom*, Farrar, Straus, & Giroux, New York, 2004.

[13] Alfred Marshall, *Principles of Economics*, p.271, Macmillan, London, 8th Bd., 1890.

[14] 固特异公司留在了阿克伦，但1990年后，阿克伦停止生产个人汽车的轮胎。

[15] 参见Johan P. Murmann, *Knowledge and Competitive Advantage*, Cambridge University Press, Cambridge, UK, 2003.

[16] 经济学家的观点参见Dosi, "Sources, Procedures, and Microeconomic Effects of Innovation," *J.Econ. Literature*, XXVI, 1120-1171, 1988; and R. Nelson and S. Winter, op. cit.

第9章　进化机制

[1]　关于早期半导体的完整的历史描述，参见Aitken, 1976, 1985.

[2]　这并不表明所有的新技术的内容都是预先存在的。原子弹爆炸的实现是通过分离方法将以前不存在的可分裂的放射性同位素U235与化学性质相似的U238分离。但是这些方法是从已有的方法当中组合而来的：包括离心分离、电磁分离、气体屏障和液体热扩散等。所以当我说技术是从已有的技术中构建而来时，我想说的是：技术是从此前技术中建构而来的，或者是从那些从已有技术中迁移而来的一两个元素中建构而来的。

[3]　Lan McNeil, "Basic Tools, Devices, and Mechanisms," in *An Encyclopedia of the History of Technology*, McNeil, ed., Routledge, London, 1990.

[4]　Ogburn, p.104.

[5]　Schumpeter, 1942, p. 82-85.

[6]　阿瑟和波拉克

[7]　Richard Lenski, etal., "The evolutionary origin of complex features, "*Nature*, 423, 139-143, 2003.

[8]　波拉克和我发现这些“沙堆”式雪崩遵循幂定律，技术性地说，就是我们的技术系统正处于自组织临界状态。

[9]　另一个更重要的来源是基因以及基因组复制。下文雅各布的话引自*The Possible and the Actual*, Pantheon, New York, 1982, p. 30.

第10章　技术进化所引发的经济进化

[1]　定义来自Dictionary.com. Word Net®3.0, Princeton University, 2008.

[2]　关于工业革命的说明，参见David S. Landes, *The Unbound Prometheus*, Cambridge University Press, Cambridge, UK, 1969; Joel Mokyr, *The Lever of Riches*, Oxford University Press, New York, 1990; and T. S. Ashton, *The Industrial Revolution*, Oxford University Press, New York, 1968.

[3]　M. E. Rose, "Social Change and the Industrial Revolution," in *The Economic History of Britain since 1700*, Vol.1, R. Floud andD. McCloskey, eds., Cambridge University Press, Cambridge, UK, 1981. See also P. WJ. Bartrip and S. B. Burman, *The Wounded Soldiers of Industry*, Clarendon Press, Oxford, UK, 1983.

[4]　M. Chase, Early Trade Unionism, Ashgate, Aldershot, UK, 2000; and W. H. Fraser, *A History of British Trade Unionism 1700-1998*, Macmillan, London, 1999.

[5]　Landes, op. cit., p. 43.

[6]　关于技术能够决定未来的经济以及未来的社会关系的观点被称为技术决定论。马克思经常因此被指责。“手推磨产生的是封建领主的社会；蒸汽磨产生的是工业资本家的社会。”罗森伯格令人信服地指出马克思太敏感了，因此并不算是一个决定论者。

[7]　弗里德利希·拉普过去曾提到过这一点。参见Paul T . Durbin, ed., *Philosophy of Technology*, Kluwer Academic Publishers, Norwell, MA, 1989.

[8]　Schumpeter, 1912.

[9]　Schumpeter, 1942, p. 82-85.

第11章　我们的立场是什么

[1] W. B. Arthur, "Complexity and the Economy," *Science* 284, 107-9, April 2, 1999.

[2] 心理学家罗伯特·约翰森（Robert Johnson）说："看起来进化的目的现在被替换为一种完美的想象和完整性或整体性的概念。完美是指某种纯净的，没有瑕疵、污点或者可疑点的东西。整体性包括黑暗，但是它与光亮共同组合于一个无以伦比的真实与整体中。这是一项令人惊叹的任务。摆在我们面前的问题是，人类是否具备这种努力和增长的能力。但是无论是否已经准备好，我们都已经上路了。"R. A. Johnson, *He: Understanding Masculine Psychology*, Harper and Row, New York, 1989, p. 64.

[3] R. Venturi, *Complexity and Contradiction in Architecture*, Museum of Modern Art, New York, 1966, p.16.

[4] Martin Heidegger, *Discourse on Thinking*, John M. Anderson and E.Hans Freund, trans.. Harper & Row, 1966.

[5] Anderson and Heidegger, introduction to Heidegger, p. xxix.

[6] Eullul, 1980, p.125.

[7] Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance*, HarperCollins, New York, 1974.

译者后记

想要认识技术的本质属性，就一定要追问“技术是怎样来的”，“是怎样的”等本体论问题。《技术的本质》这本书可以说正是从这个角度回答了我们想要知道的东西。这也是我们决定翻译这本书的初衷。

在这里，技术思想家布莱恩·阿瑟从“技术黑箱”的内部探究技术的本质，形成了一套完整的技术本质及其进化的理论。这个理论框架揭示了技术内部的三个基本原理及其逻辑结构的生成机制，以及这个结构在其最深的本质上展现的进化的共性。在方法论上，尝试通过打开“技术黑箱”来“看”技术所显现的技术本质及其进化机制。整个理论使得以往关于技术本质的技术哲学、工程哲学、设计哲学、经济学、社会学、甚至科学上的众多观点在这个框架下得到极大程度的统一。

那么通过这本书，我们能够得到哪些启示呢？首先，技术体的概念对于解释不同文化背景下、不同科学技术发展道路上的国家和地区的技术发展程度之所以不同具有深刻的启示作用。其次，对于创新和创造的本源的揭示，揭开了那层神秘的面纱，直指其根本：没有科学技术长期的积累和传承，是很难在新颖性上有所建树的。我们认为，本书呈现的是关于技术理论的一个全貌，对技术及其进化的本质作了直言不讳的宣告：技术作为一个“体”的产生和进化对揭示技术的有机性给出了强有力的证据，随后是对技术和经济的关系的全新阐释，最后对“我们应该怎样看待技术”的追问因此显得深刻有力。

从“技术的组合进化理论”中，我们不时地看到熟悉的身影，达尔文、海德格尔、熊彼特、埃吕尔、库恩等，以及许多科学家、技术专家、社会学家、技术史家、经济学家，事实上，很少有人将这些人的观点和方法全部都统一到一个关于技术的理论框架之下。是布莱恩将技术内部的解剖结构首先在技术的一般层面上有条理地呈现出来；进而把这种解剖结构放回到真实世界（real world）当中，从而改变以往关于技术的静态逻辑分析，“看”到技术真实的动态变化，发现这是一个技术的自创生的进化事件。技术由人类创造出来，又基于自然最原初的现象，但却开始了疏离“人类与自然”的进化之旅。笔者认为这是一个令人惊奇的结论。

本书内容涉及数十种具体技术，作者对它们逐一进行了解构和历史

进化的追述，这也构成了翻译过程中最困难的部分。因为作者希望这是一本用“平实”的语言讲故事的书，是一本面向“大众”而不是专家的书。如何既准确呈现大量关于科学技术的专业叙述，又能反映出作者意在进行松、浅显甚至幽默的表达也颇费思虑。我们在翻译过程中秉承了这样的一个宗旨，尽量用生活化的语言将晦涩的专业术语所要表达的意思表达出来，希望结果能够达成作者的本意。

能够翻译本书首先要感谢远德玉教授，是他以敏锐的学术感知力洞察到这本书的独特之处。2009年，在这本书的英文原版刚刚出版之际，远德玉教授就推荐给我们进行阅读，随后又提出，希望我们能够将它翻译出来，并在翻译过程中不遗余力地贡献出所有见解。此外，王晓航博士对本书译稿进行了细致而全面的校对，提出许多宝贵的意见和建议，其工作细致而艰苦，在这里我们也对他的无私帮助致以深深的谢意。同时也要感谢罗玲玲、马会端、朱春艳、董雪林、吴俊杰等同仁在翻译过程中提供的帮助，以及我们的家人、朋友在背后默默的支持。正因为有他们，我们在漫长孤独的学术旅途上才有归宿感、安定感，以及继续探索下去的勇气。而本书得以出版，更要感激湛庐文化的大力支持，以及简学老师的不懈努力和辛苦工作，再次向他们表示衷心的感谢！

最后，限于译者水平，错漏之处在所难免，敬请各位读者朋友批评指正。

曹东溟　王健

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读APP：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当做阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”APP。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读APP玩转指南

湛庐阅读APP结构图：



三步玩转湛庐阅读APP：



APP获取方式：

安卓用户前往各大应用市场、苹果用户前往APP Store

直接下载“湛庐阅读”APP，与最聪明的人共同进化！

使用APP扫一扫功能，
遇见书里书外更大的世界！



快速了解本书内容，
湛庐千册图书一键购买！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

延伸阅读

《复杂经济学》

- ◎ 湛庐文化圣塔菲书系的首部著作！
- ◎ 布莱恩·阿瑟继《技术的本质》之后，又一部重磅作品！
- ◎ 经济学的新古典主义时代已经结束，这本书是“复杂经济学”的奠基之作！





《同步：秩序如何从混沌中涌现》

- ◎ “复杂性理论”开创性研究者斯蒂芬·斯托加茨重磅新作！
- ◎ 互联网时代最重要的观念之一“小世界网络”的提出者！
- ◎ 电子科技大学教授周涛，诺贝尔奖获得者菲利普·安德森，牛津大学教授罗伯特·梅，美国海军研究实验室卢·佩科拉等联袂推荐！





《重塑：信息经济的结构》

- ◎ “新经济丛书”首部著作，由信息经济先行者张翼成，复杂网络研究青年学者吕琳媛、周涛合力打造，是奠定新经济体系的思想基石！
- ◎ 阿里巴巴集团学术委员会主席曾鸣重磅推荐！信息社会50人论坛理事、阿里研究院高级顾问梁春晓，财讯传媒集团首席战略官、网络智酷总顾问段永朝亲笔作序，鼎力推荐！





《超级合作者》

◎ 新时代的达尔文、明星科学家马丁·诺瓦克前沿研究之集合，创立第3进化原则，破解人类合作之谜。

◎ 洞悉人类社会与行为的里程碑式著作。





(1) 本书中以“1”标注的注释内容详见书后的注释，以“①”标注的注释内容详见每页的页下注。——编者注

(2) 指由杰斐逊总统发起的美国国内首次横越大陆，西抵太平洋沿岸的往返考察活动。该活动由刘易斯上尉和克拉克少尉担任领队。——编者注

(3) 阿瑟在这里所用的“descent”一词，也正是达尔文写《人类的由来》（*The Descent of Man*）的书名所用的词，译者参考了商务印书馆潘光旦、胡寿文译本的译注，“descent”一词在书名中译为“由来”，而在正文中译作“诞降”，前辈的深思熟虑值得借鉴，后辈不敢掠美，谨以为识。——译者注

(4) 生态学中环境的概念，亦称栖息地，指生物生活的空间和其全部生态因子的总和。——编者注

(5) 这种诗的结构是：每3行必须要包含17个音节，分别是第1行有5个音节，第2行有7个音节，第3行有5个音节。——译者注

(6) 著名经济学家，第十届诺贝尔经济学奖获得者，主要研究领域为经济组织的管理行为和决策行为。——编者注

(7) 天文学中指太阳黑子周围的较淡的部分。——译者注

(8) 拇指法则，即一种经验性、直觉性的简单原则。——编者注

(9) 指法国顶级的厨艺水平。——编者注

(10) 本轮（epicycle）是天文学中的术语，在“地心说”时代，古人原本认为其他行星按照同一方向围绕地球转，但是又发现有的行星在逆行，为了解释这一现象，又提出了本轮概念。——译者注

(11)　这里的基因工程主要是指其医学应用方面：其在本质上是利用基因去生产或控制人类健康所必需的蛋白质。[2]

(12)　可参考湛庐文化2018年5月出版的阿瑟的新书《复杂经济学》。——编者注

CHEERS
湛庐

模型思维

What You Need to Know to Make Data Work for You

超过100万用户的
模型思维课
首创者佩奇教你
如何聪明决策

THE MODEL THINKER


圣塔菲书系 Santa Fe Institute Series


浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

[美] 斯科特·佩奇 著
Scott Page

贾拥民 译

版权信息

本书纸版由浙江人民出版社于2019年12月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

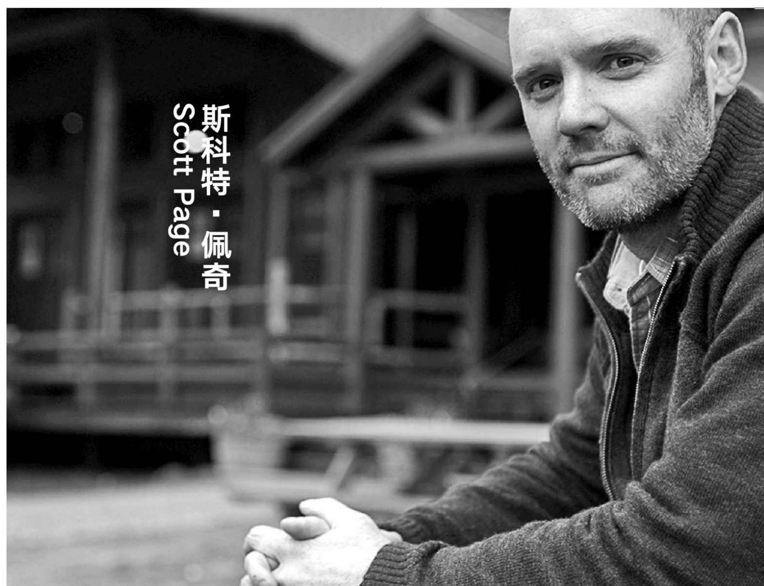
书名：模型思维

著者：斯科特·佩奇

电子书定价：134.99元

THE MODEL THINKER：What You Need to Know to Make Data Work for You By Scott E. Page.

Copyright © 2018 by Scott E. Page. All rights reserved



广受欢迎的“模型思维课”主讲人

密歇根大学复杂性研究中心“掌门人”

圣塔菲研究所外聘研究员

研究复杂性与多样性的专家



斯科特·佩奇于1985年获得密歇根大学安阿伯分校数学学士学位，1988年获得威斯康星大学麦迪逊分校数学硕士学位，1990年获得西北大学凯洛格商学院管理经济学硕士学位，1993年获得西北大学凯洛格商学院管理经济学和决策科学博士学位。

佩奇以对社会科学的多样性和复杂性的研究和建模而闻名。具体研究方向包括路径依赖、文化、集体智慧、适应和社会生活的计算模型。

研究领域涉及多个学科，包括经济学、政治学、计算机科学、管理学、物理学、公共卫生、地理学、城市规划、工程学和历史学。

他曾多次在高中、大学、公司、非营利组织以及政府演讲，介绍他关于多样性和复杂性的研究。也曾经为国际货币基金组织、美国教育部、福特汽车公司、奔驰汽车公司等提供咨询。

佩奇获得了多项奖金，包括2002年的IGERT奖和2001—2006年的生物复杂性项目SLUCE奖，以及2013年的古根海姆奖（Guggenheim Fellowship）。他还曾多次获得加州理工学院、西北大学和密歇根大学颁发的杰出教学奖，这些大学对他多年来在复杂性和多样性方面的教学成果给予了高度认可。佩奇于2011年当选美国艺术与科学学院院士。

广受欢迎的“模型思维课”主讲人

密歇根大学前校长玛丽·苏·科尔曼（Mary Sue Coleman）说：“我们的教师渴望与全世界分享他们的知识，我们的学生对于可以体验全新的教学方式也同样激动。”斯科特·佩奇对此也积极响应，他在Coursera平台上线了“模型思维课”，该课程包括一百多个视频和阅读资料。自课程上线以来，有超过5万名学生注册该课程，超过120万人次观看了课程视频，受到了来自世界各地学生的好评。



佩奇教授鼓励对课程好奇的人们注册学习，并亲身参与其中。他

说："这是一个很好的机会，让我们的校友和想要就读密歇根大学的学生们体验一下什么是密歇根大学的教育方式。"

在这门课程里，佩奇讲授了理解和应用模型如何帮助人们做出更好的决策。有证据表明，具备模型思维的人要比没有这种思维的人更优秀，而且能够运用多种模型思考并解决问题的人要比只运用一种模型思考的人更优秀。

佩奇在课程里还着重介绍了几种模型，展示出多样性对创新的重要性。并具体讲解了拥有多样性视角、启发式的群体是怎么比个人表现更好的。

佩奇的课程引用大量的例子，内容生动有趣。他曾在模型思维课程里讲过一个观点：在离散状态马尔科夫（数学模型）过程中，如果把人生看成努力和不努力两个状态，只要状态转移矩阵确定了，长期来看，在每个状态下所停留的时间比例也就都确定了。如果人生的动力源泉是固定的，努力的百分比就是固定的，那么短期内努力或者不努力并不会有什么影响。也就是说，问题的根本不在于你的状态，而在于源动力！所以，在瓶颈期遇到困难实在不想努力的话，多去找一找自己的源动力，想想当初为什么出发。

密歇根大学复杂性研究中心"掌门人"圣塔菲研究所外聘研究员



斯科特·佩奇于2002—2009年担任密歇根大学复杂性研究中心副主任，于2010—2015年担任主任。复杂性研究中心成立于1999年，其创始成员是一个现已成为传奇的研究小组——巴赫小组（BACH Group）。巴赫小组始于20世纪80年代，成员包括来自不同学科的研究人员，他们都对各种复杂的适应性系统感兴趣。





巴赫小组最初的成员包括美国数学家亚瑟·伯克斯（Arthur Burks）、遗传算法之父约翰·霍兰德（John Holland）等人。美国进化生物学家威廉·汉密尔顿（William Hamilton）、认知科学家侯世达、认知科学家梅勒妮·米歇尔（Melanie Mitchell）也是巴赫小组的成员。现在，巴赫小组由政治学家罗伯特·阿克塞尔罗德（Bob Axelrod）、物理学家马克·纽曼（Mark Newman）、数学家和公共政策学家卡尔·西蒙（Carl Simon）和斯科特·佩奇等人组成。其中，马克·纽曼、卡尔·西蒙也是复杂性研究中心的成员。

佩奇于1999年被圣塔菲研究所聘为研究员，开始了在圣塔菲研究所十几年的研究生涯，主要研究方向依然是复杂性和多样性。

作者演讲洽谈，请联系
speech@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS 特别制作

献给

迈克尔·科恩

不应否认，任何理论的终极目标都是尽可能让不可简化的基本元素变得更加简单且更少，但也不能放弃对任何一个单一经验数据的充分阐释。

阿尔伯特·爱因斯坦



扫码下载“湛庐阅读”APP，
搜索“模型思维”，
获取更多精彩内容。

序言　这本书是怎样写成的

对我来说，成功意味着我在这个世界上的有效性：我能够把我的思想和价值观带给这个世界，我能够以积极的方式改变它。

汤亭亭

这本书源于我与美国复杂系统专家迈克尔·科恩（Michael　Cohen）一次偶然的见面。那是在2005年，密歇根大学西厅旁边购物中心的花园。作为一名学者，迈克尔素以慷慨大方而闻名，他在这次见面时说的一番话，彻底改变了我的教学生涯。迈克尔眼中闪烁着光芒，他说："斯科特，我曾经根据查尔斯·拉夫（Charles　Lave）和詹姆斯·马奇（James　March）写的一本书，开设过一门名为'社会科学建模导论'的课程。你应该重开这门课程，它需要你。"

它需要我？我有点困惑。回到办公室后，我马上找出了那门课的课程大纲。看了之后，我发现迈克尔错了，不是这门课程需要我，而是我需要它。我一直在寻找这样一门课程：既能够向学生介绍复杂系统领域的知识，同时又不脱离他们的日常生活和未来的职业规划。通过讲授一门关于模型的课程，我可以向学生展示各种相关的工具和思想，提高他们推理、解释、设计、沟通、行动、预测和探索的能力。

开设这样一门课程的基本动机是，我们必须利用多种多样的模型去应对复杂性。学了一个学期之后，学生们在看待这个世界时，就不会拘泥于某个特定的角度，相反，他们将透过多种不同的视角去观察世界。他们将站在有许多扇窗户的房子中，拥有看向多个方向的能力。我的学生应该能够更好地应对他们所面对的复杂挑战：改善教育，减少贫困，实现可持续增长，在人工智能时代找到有意义的工作，管理好资源，设计出强大而稳健的金融、经济和政治体系。

第二年秋天，我真的重新开设了这门课程。我本来打算将这门课程命名为"32种使你变成天才的模型"，但是密歇根大学的传统不允许这个有"王婆卖瓜"之嫌的名字。于是，我沿用了迈克尔的课程名称"建模导论"。事实早已证明，查尔斯·拉夫和詹姆斯·马奇的书是非常不错的入门读物。不过，在过去的几十年中，建模技术已经取得了巨大的进步。我需要对这门课程进行更新，以便将长尾分布（long-tailed

distributions）、网络模型、崎岖景观模型（rugged-landscape model）和随机游走模型（random walk model）等全都包括进来。我需要一本讨论复杂性的书。

于是，我开始了写作，但这并非易事。刚开始的两年，写作进展得非常缓慢。后来，在一个春天，我又一次遇到了迈克尔，这次是在西厅的拱门那里。我自己一直对这门课程有所疑问，尽管那时课程已经吸引20多名学生了。对于本科生来说，这些模型是否过于抽象了？我是不是应该针对不同的问题或政策领域开设不同的课程呢？面对这些问题，迈克尔笑了笑，他指出任何值得做的事情都必定会招致质疑。告别时，迈克尔再一次强调，帮助人们清晰地进行思考是非常重要且非常有价值的。他叮嘱我不要放弃。

2012年秋季学期，这门课程又发生了根本性变化。密歇根大学副教务长玛莎·波拉克（Martha Pollack）邀请我加入在线开放课程体系，开设一门在线课程，也就是现在所说的慕课（MOOC）。于是，借助于一台联网的电脑，一个29美元的摄像头和一个90美元的麦克风，“模型思维”（Model Thinking）这门在线课程正式诞生了。

在来自密歇根大学、斯坦福大学以及Coursera的无数师友的帮助下，我将上课用的讲义重新改编成了适合在线课程的形式，包括将每个主题划分为若干模块，并删除了所有受版权保护的资料。在这里，我要对这些热心人士表示感谢，尤其是汤姆·希基（Tom Hickey），他总是随叫随到，帮我解决了很多疑难问题。为了保证效果，我一遍又一遍地重新录制课程。感谢我的狗Bounder，它是我忠实的听众。

“模型思维”第一次上线时就吸引了6万名学生。到现在，这门课程的注册学生人数已经接近100万了。在线课程如此受欢迎，以至于我暂时放弃了写作本书的计划。我当时认为也许没有必要写一本书。但是很快，我的电子邮件收件箱就被学生们要求得到一本可以作为在线课程补充书籍的请求塞满了。在迈克尔·科恩不幸因癌症去世后，我更加觉得自己必须尽快完成这本书，于是我重新摊开了书稿。第1稿，第2稿……第9稿……现在，你们看到的是第30稿。

写一本书需要大量的时间和开阔的空间。美国著名现代诗人华莱士·史蒂文斯（Wallace Stevens）曾经这样写道：“也许，真理取决于在湖边散步的时间。”就我自己写作这本书的经历来说，思考大大得益于数百次横穿怀南斯湖（Winans Lake）的游泳，每到夏天，我和我的家人总是在那里过周末。

在整个写作过程中，我和我生命中的挚爱珍娜·贝德纳（Jenna Bednar）、我们的儿子奥里（Orrie）和库珀（Cooper），以及我们养的3只爱犬Bounder、Oda和Hildy一起开怀大笑、享受生命、面对挑战、把握机遇。在那里，奥里订正了本书倒数第二稿中的数学错误，珍娜花了两个星期阅读了全部书稿并修正了数百个有问题的地方：字体和语法上有错误的、思路不清晰的、逻辑有缺陷的、举例混乱的……事实上，与我撰写的所有文字作品一样，对这本书的一个更加准确的描述是：斯科特·佩奇写出了第一稿，最后定稿则由珍娜·贝德纳完成。

在写作本书的过程中，我看着我的孩子们从青春期过渡到了成人阶段。我们一家人也吃掉了太多的石锅拌饭、意大利面和燕麦片巧克力饼干。我们用锯子和修枝剪，修剪了无数棵大大小小的树木，补好了篱笆上的几十个破洞；花费了很多时间和精力，试图让地下室和车库变得整齐一些，但最终却总是无功而返。我们也一年又一年地盼望湖上的冰厚到可以滑冰……

从草拟大纲到提交定稿，用了整整7年的时间，怀南斯湖冬天结冰、春天又化开，前前后后也有7次了。然而，有的冰冻期比其他冰冻期更加难挨。在写作本书的过程中，我的母亲玛丽莲·坦博·佩奇（Marilyn Tamboer Page）不幸因心脏病突发去世，当时她正在散步，那是她每天的例行事务之一。有些深洞永远无法填满，它们会提醒我们珍惜生命所提供的宝贵机会。

现在，这本书终于已经完成了。奥里已经上了大学，库珀明年也要去上大学了。如果您，我亲爱的读者，认为本书中的模型和思想是有用的、有创造性的，而且能够将它们应用于现实世界，并以积极的方式去改变世界，那么我将这些资料组织成书的努力就得到了最大的回报。如果有一天，当我在某个教授或研究生的办公室里（最有可能是在美国中西部的某所大学里）浏览书架时，在拉夫和马奇的书旁边发现了我的这本书，那么我将会觉得非常幸福。

目 录

序言　这本书是怎样写成的
Part 1　为什么需要模型思维
01　做一个多模型思考者
02　模型的7大用途
03　多模型思维
04　对人类行为者建模
Part 2　模型思维
05　正态分布
06　幂律分布
07　线性模型
08　非线性模型
09　与价值和权力有关的模型
10　网络模型
11　广播模型、扩散模型和传染模型
12　熵：对不确定性建模
13　随机游走
14　路径依赖模型
15　局部互动模型
16　李雅普诺夫函数与均衡
17　马尔可夫模型
18　系统动力学模型
19　基于阈值的模型
20　空间竞争模型与享受竞争模型
21　博弈论模型
22　合作模型
23　与集体行动有关的问题
24　与机制设计有关的模型
25　信号模型
26　学习模型
27　多臂老虎机问题
28　崎岖景观模型
结语　像芒格一样智慧地思考——多模型思维的实际应用
译者后记
注释

参考文献

The Model Thinker

What You Need to Know to Make Data Work for You

Part 1 为什么需要模型思维

多模型思维要求掌握多个模型，但是我们并不需要懂得非常多的模型，只要知道每个模型都可以有多种应用场景。要成为一个多模型思考者，需要的不仅仅是数学能力，更需要的是创造力。

01 做一个多模型思考者

要想成为一个有智慧的人，你必须拥有多个模型。而且，你必须将你的经验，无论是间接的，还是直接的，都放到构成这些模型的网格上。

查理·芒格（Charlie Munger）

这是一本关于模型的书。我在书中用简洁的语言描述了几十个模型，并解释该如何应用它们。模型是用数学公式和图表展现的形式化结构，它能够帮助我们理解世界。掌握各种模型，可以提高你的推理、解释、设计、沟通、行动、预测和探索的能力。

本书提倡多模型思维方法，应用模型集合理解复杂现象。本书的核心思想是：多模型思维能够通过一系列不同的逻辑框架“生成”智慧。不同的模型可以将不同的力量分别突显出来，它们提供的见解和含义相互重叠并交织在一起。利用多模型框架，我们就能实现对世界丰富且细致入微的理解。本书还包括了一些正式的论证，阐述了如何对现实世界应用多模型框架。

本书非常实用。多模型思维具有十分重要的实用价值。运用这种思维方式，你就能更好地理解复杂现象，就能更好地推理。你将会在职业生涯、社区活动和个人生活中表现出更小的差距，做出更加合理的决策。是的，你甚至还可能会变得更有智慧。

25年前，像本书这样讲解模型的著作主要是供教授们和研究生们研究商业、政策和社会科学所用的，金融分析师、精算师和情报界人士也是潜在的读者。这些人都是应用模型的人，他们也是与大型数据库关系最密切的人，这并不是偶然。不过到了今天，关于模型的书已经拥有了更多的读者：广大的知识工作者们。由于大数据的兴起，他们现在已经把模型作为日常生活的一部分了。

如今，用模型组织和解释数据的能力，已经成了商业策略家、城市规划师、经济学家、医疗专家、工程师、精算师和环境科学家等专业人士的“核心竞争力”。任何人，只要想分析数据、制订业务发展策略、分配资源、设计产品、起草协议就必须应用模型，哪怕是做出一个简单招聘决策，也要运用模型思维。因此，掌握本书的内容，特别是那些涉及

创新、预测、数据处理、学习和市场准入时间选择的模型，对许多人都有非常重要的实际价值。

使用模型来思考能够带给你的，远远不仅仅是工作绩效的提高。它还会使你成为一个更优秀的人，让你拥有更强的思考能力。你将更擅长评估层出不穷的经济事件和政治事件，更能识别出自己和他人推理中的逻辑错误。有了这种思维方式，你将懂得辨识什么时候意识形态取代了理性思考，并对各种各样的政策建议有更丰富、更有层次的洞见，无论是扩建城市绿地的建议，还是强制药物检测的规定。

所有这些好处都来自与多种多样模型的“亲密接触”，幸运的是，我们用不着一下子掌握千百种模型，而只需先掌握几十种就足够了。本书给出的这些模型就为你提供了一个很好的出发点。它们来自多门学科，其中包括许多人耳熟能详的囚徒困境博弈模型，逐底竞争（Race to the Bottom）和关于传染病传播的SIR模型，等等。所有这些模型都有一个共同的形式：它们都假设一些实体，通常是人或组织，并描述他（它）们是如何相互作用的。

本书所讨论的模型可以分为三类：对世界进行简化的模型、用数学概率来类比的模型以及人工构造的探索性模型。无论哪一种形式，模型都必须是易处理的。模型必须足够简单，以便让我们可以在模型中应用逻辑推理。例如，我们讨论了一种传染病模型，这个模型由易感者、感染者和痊愈者组成，可以给出传染病的发生概率。利用这个模型，我们可以推导出一个传染阈值，也就是一个临界点，超过这个临界点，传染病就会传播。我们还可以确定，为了阻止传染病传播，需要接种疫苗人数的比例。

尽管单个模型本身可能就已经相当强大了，但是一组模型可以实现更多的功能。在拥有多个模型的情况下，我们能够避免每个模型本身所固有的局限性。多模型方法能够消除每个单个模型的盲点。基于单一模型的政治选择可能忽略了世界的一些重要特征，如收入差距、身份多样性以及与其他系统的相互依赖关系。[1]有了多个模型，我们可以达成对多个流程的逻辑推理，可以观察不同因果过程是如何重叠和相互作用的，也拥有了理解经济、政治和社会世界复杂性的可能。而且，我们在这样做的时候并不需要放弃严谨性，因为模型思维能够确保逻辑的一致性。由此，推理将建立在扎实的证据基础之上，因为模型需要用数据检验、改进和精炼。总而言之，当我们的思维得以在多个逻辑上一致、处在通过了经验验证的框架中时，我们更有可能做出明智的选择。

大数据时代的模型

在当今这个大数据时代，像本书这样一本讨论模型的书可能看上去有些不合时宜。现在，数据正以前所未有的维度和粒度急速地涌现出来。过去，消费者的购买数据只能以每月汇总表的形式打印出来，而现在却可以与空间、时间信息及消费者“标签”一起实时传输。学生的学习成绩数据，现在也包括每一份作业、每一篇论文、每一次测验和考试的分数，而不再仅仅是一个期末总成绩了。过去，农场工人也许只能在每月一次的农场会议上提出土壤过于干燥的问题，而现在，他们却能够用拖拉机自动传输以平方米为单位的关于土壤肥力和水分含量的实时数据了。投资公司要跟踪数千只股票的数十种比率和趋势，并使用自然语言处理工具来解析文档。医生则可以随时提取包括相关遗传标记在内的患者记录。

仅仅在25年以前，大多数人获得的知识只能来自书架上的几本书。也许你工作的地方有一个小型图书馆，或者你家里有全系列的百科全书和几十本参考书。学术界、政府和私营部门的研究者则可以利用大型图书馆的馆藏资料，但是他们也经常不得不亲身前往查阅。就在20世纪末21世纪初，为了获得必要的信息，学者们仍然不得不在卡片目录室、缩微胶片阅览室、图书馆书架以及私人收藏家的“宝库”之间来往穿梭。

现在，这一切都发生了颠覆性的变化。几个世纪以来一直受到纸张束缚的知识内容，今天已经以数据包的形式在“空中”自由流动了。关于此时此地的实时信息也是如此。以前，新闻是刊载在报纸上的，最高以每天一次的频率送到我们手上；而现在，新闻却是以连续的数字流形式流入我们的个人设备。股票价格、体育赛事比分、关于政治经济事件和文化事件的新闻，全都可以实时查询、实时访问。

然而，无论数据给我们留下的印象如何深刻，它都不是灵丹妙药。我们也许可以通过数据了解到已经发生了什么和正在发生什么，但是，由于现代世界是高度复杂的，我们可能很难能理解为什么会发生这种情况。更何况，经验事实本身也可能是误导性的。例如，关于计件工资制的统计数据往往会显示，工人每生产一件产品获得的报酬越高，他们生产的产品就会越少。对此，用一个薪酬取决于工作条件的模型可以很好地解释相关数据。如果工作条件很差，导致很难生产出产品，那么每单位产品的工资可能很高；如果工作条件很好，那么每单位产品的工资就可能会很低。因此，并不是更高的计件工资导致了更低的生产率，而是

更加糟糕的工作条件导致了这种结果。[2]

此外，我们社会中的大多数数据，也就是关于经济、社会和政治现象的数据，都只是时间长河上的瞬间或片断的记录。这种数据是不能告诉我们普遍真理的。我们的经济、社会和政治世界并不是固定不变的。在这个十年内，男孩在标准化考试中的成绩超过了女孩，但是下个十年就有可能变为女孩的成绩好于男孩。人们今天投票的原因，可能与未来几十年投票的原因截然不同。

我们需要模型，不然就无法理解计算机屏幕上不断滑过的数据流。因此，这个时代，可能恰恰因为我们拥有如此多的数据，也可以被称为多模型时代。纵观学术界、政府、商界和非营利部门，你基本上无法找到任何一个不受模型影响的研究领域，甚至可以说根据不存在不需要模型的决策领域。麦肯锡（McKinsey）和德勤（Deloitte）等咨询业巨头要通过构建模型来制订商业策略；贝莱德集团（BlackRock）和摩根大通集团（JPMorgan Chase）等金融业大公司要利用模型来选择投资，州立农业保险公司（State Farm）和美国好事达保险公司（Allstate）等公司的精算师要借助风险校正模型来给保险单定价。谷歌公司的人力资源部门要利用预测分析模型来为超过300万求职者进行评估。各大学和学院的招生人员也要建立模型，以便从成千上万的申请入学者当中选出合格的新生。

美国行政管理和预算局（Office of Management and Budget）通过构建经济模型预测税收政策的影响。华纳兄弟公司通过数据分析模型评估观众对电影的反应。亚马逊公司开发机器学习模型向消费者推荐商品。由美国国家卫生研究院（National Institutes of Health）资助的研究团队建立了人类基因组学的数学模型，用于寻找和评估癌症潜在的治疗方法。盖茨基金会使用流行病学模型设计疫苗接种策略。甚至运动队也都使用模型来预测选秀结果和交易机会，并制订比赛策略。例如，芝加哥小熊队（Chicago Cubs）之所以能够在经历了一个多世纪的失败后赢得世界职业棒球联赛的冠军，就是因为很好地利用了模型去选择球员、设计比赛策略。

对于使用模型的人来说，模型思维的兴起还有一个更简单的解释：模型能够让我们变得更聪明。如果没有模型，人们就会受到各种认知偏差的影响：我们会对近期发生的事件赋予过高的权重、会根据“合理程度”分配概率、会忽略各种基本比率。如果没有模型，我们处理数据的能力就会受到极大的限制。有了模型，我们就能澄清相关假设且更有逻辑

地进行思考，还可以利用大数据来拟合、校准、检验因果关系与相关性。总之，有了模型，我们的思考会更有效。有证明表明，如果让模型与人面对面直接“竞争”，模型将会胜出。[3]

为什么需要多模型

在本书中，我们主张在给定情况下不仅使用一个模型，而要使用多个模型。多模型方法背后的原理基于这样一个古老的思想，那就是“管中窥豹需多管齐下”。这个思想至少可以追溯至亚里士多德，他强调了将许多人的优点集中起来这个做法的价值。呈现视角和观点的多样性，也是美国历史上“名著运动”（great-books movement）背后的一大动力。在这个运动中涌现出来的《伟大的思想：西方世界名著中伟大的思想观念合集》（*The Great Ideas*：*A Syntopicon of Great Books of the Western World*）一书，就收集了102个重要的可永世流传的思想。

现在，这种方法也在汤亭亭所著的《女勇士》（*The Women Warrior*）一书中得到了回响，她这样写道：“我已经学会了如何让我的思想变得博大；因为宇宙很大，所以给悖论留下了存在的余地。”这种方法也构成了现实的商业和政治世界有实际意义的行动基础。最近的一些论著指出，如果我们想要理解国际关系，就不能只将世界建模为一组具有明确目标的自利国家，也不能只将世界建模为跨国公司和政府间组织之间的联系枢纽，而应该把世界同时建模为这两者。[4]

尽管多模型方法看上去似乎很平常，但请注意，它其实是与我们讲授模型和构建模型的传统方法相悖的。传统的方法，那些在高中时老师教授的方法，依赖一对一的逻辑，也就是说一个问题需要一个模型。比如，老师会告诉我们，在这种情况下，我们应该运用牛顿第一定律；在那种情况下，我们应该运用牛顿第二定律；在第三种情况下，则应该运用牛顿第三定律。又或者，在这里，我们应该使用复制因子方程（replicator equation）来说明下一期兔子种群的大小。在这种传统的方法中，目标是确定一个适当的模型并正确应用这个模型。而多模型思维所要挑战的，恰恰正是这种传统方法。多模型方法主张尝试多个模型。如果你在九年级时就使用过多模型思维，你可能会被阻止，但是现在使用多模型思维，你将会取得很大进步。

大部分学术论文也遵循传统的一对一的方法，尽管有时它们是在使用单一的模型去解释复杂的现象。例如，有人声称，在美国2016年选举

中投票给特朗普的那些人，都是经济上的失败者。又或者，小学二年级时老师的素质决定了孩子长大成人后能够取得经济成就的大小。[5]不过，近年来，一系列畅销的非虚构作品的诊断，使这种基于单个模型的传统思维方式的弊端呈现在人们面前：教育成功只取决于毅力；资本集中导致不平等；糖消耗导致民众健康状况不佳……这些单个模型中的每一个都可能是正确的，但没有一个是全面的。面对各种复杂的挑战，创造一个包容更广泛教育成就的世界，我们需要的不是单个模型，而是多个模型构成的格栅。

通过学习本书中的模型，你就可以着手构建自己的格栅模型。这些模型来自多个学科，涉及各种现象，例如收入不平等的原因、权力的分配、传染病和流行风尚的传播、社会动乱的前置条件、合作的发展、秩序的涌现，以及城市和互联网的结构等。

模型的假设和结构各不相同。有些模型描述了少量理性的、自私的行为主体之间的互动，有些模型则描述了大量的遵循规则的利他主义者的行为。一些模型描述了均衡过程，还有一些模型讨论路径依赖性和复杂性。这些模型的用途也各不相同。一些模型是用来帮助预测和解释的，一些模型是用来指导行动、推动设计或促进沟通的，还有一些模型则创造了有待我们去探索的虚拟世界。

所有模型都有三个共同特征。第一，它们都要简化，剥离不必要的细节，抽象掉若干现实世界中的因素，或者需要从头重新创造。第二，它们都是形式化的，要给出精确的定义。模型通常要使用数学公式，而不是文字。模型可以将信念表示为世界状态的概率分布，可以将偏好表示为各备选项之间的排序。通过简化和精确化，模型可以创造易于处理的空间，我们可以在这些空间上进行逻辑推理、提出假说、设计解决方案和拟合数据。模型创建了我们能够以符合逻辑的方式进行思考的结构。正如维特根斯坦在《逻辑哲学论》（*Tractatus Logico-Philosophicus*）一书中所写的："逻辑本身就能解决问题，我们所要做的，就是观察它是如何做到的。"是的，逻辑有助于解释、预测、沟通和设计。但是，逻辑也不是没有代价的，这就导致模型的第三个共同特征是：所有模型都是错误的，正如统计学大师乔治·博克斯（George Box）所指出的那样。[6]所有模型概莫能外，即使是牛顿提出的那些定律和法则，也只是在特定的条件下成立。所有模型都是错误的，还因为它们都是简化的，它们省略掉了细节。通过同时考虑多个模型，我们可以实现多个可能情况的交叉，从而克服单个模型因严格而导致的狭隘性。

只依靠单个模型其实是过于狂妄自大的表现，这种做法会导致灾难性的后果。相信只凭一个方程，就可以解释或预测复杂的现实世界现象，会使真理成为那种很有“魅力”的简洁的数学公式的牺牲品。事实上，我们永远不应指望任何一个模型能够准确预测1万年后的海平面将上升多少，甚至也不应该指望任何一个模型能够准确预测10个月后的失业率。我们需要同时利用多个模型才能理解复杂系统。政治、经济、国际关系或者大脑等复杂系统永远都在变化，时刻都会涌现出介于有序和随机之间的结构和模式。当然，根据定义，复杂现象肯定是很难解释或预测的。[7]

因此在这里，我们面临着一个严重的脱节。一方面，我们需要模型来连贯地思考。另一方面，任何只具有少数几个活动部件的单个模型都无法解释高维度的复杂现象，例如国际贸易政策中的模式、快速消费品行业的发展趋势或大脑内部的适应性反应。即便是牛顿，也无法写出一个能够解释就业水平、选举结果或犯罪率下降趋势的三变量方程。如果我们希望了解传染病的传播机制、教育成效的变化、动植物种类的多样性、人工智能对就业市场的冲击、人类活动对地球气候的影响，或者社会动乱的可能性，就必须通过多个模型去了解它们：机器学习模型、系统动力学模型、博弈论模型和基于主体的模型等。

智慧层次结构

为了论证多模型思维方式的优点，我们先从诗人和剧作家T.　S.艾略特的一个疑问入手：“我们迷失于知识中的智慧到哪里去了？我们迷失于信息中的知识到哪里去了？”在这里，我们还可以加上一句：我们迷失于数据中的信息到哪里去了？

我们可以把艾略特的这个疑问形式化为一个智慧层次结构（wisdom hierarchy），如图1-1所示。在这个智慧层次结构的最底部是数据，也就是原始的、未编码的事件、经历和现象。出生、死亡、市场交易、投票、音乐下载、降水、足球比赛，以及各种各样的（物种）发生事件等。数据既可以是一长串的0和1，也可以是时间戳，或是页面之间的链接等。数据是缺乏意义、组织或结构的。

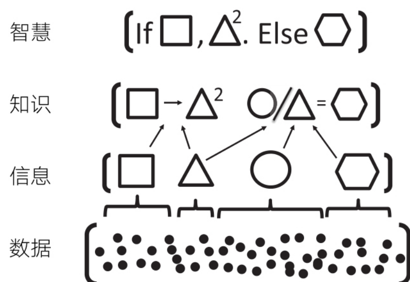


图1-1　智慧层次结构

信息用来给数据命名并将数据归入相应的类别。为了说明数据与信息之间的区别，看看这几个例子：落在你头上的雨是数据，佛蒙特州伯灵顿市和安大略湖的7月份总降水量则是信息；威斯康星州麦迪逊市国会大厦旁边周六市场上的鲜红辣椒和金黄玉米是数据，而农民的总销售额则是信息。

我们生活在一个信息极大丰富的时代。一个半世纪以前，掌握信息可以带来很高的经济和社会地位。英国小说家简·奥斯汀（Jane Austen）笔下的爱玛就曾问过，弗兰克·丘吉尔（Frank Churchill）是不是“一个拥有着很多信息的年轻人”。(1)如果放到今天，她肯定不会在意这个问题。如果穿越到现在，那么弗兰克·丘吉尔会和其他人一样有一部智能手机，问题只在于他有没有能力很好地利用这些信息。正如陀思妥耶夫斯基在《罪与罚》一书中所写的那样：“他们说，我们已经得知了事实。但事实不是一切，至少有一半的分歧就出在人们怎样利用事实上！”

柏拉图将知识定义为合理的真实信念。更现代的定义则认为知识就是对相关关系、因果关系和逻辑关系的理解。知识组织了信息，呈现为模型的形式。市场竞争的经济学模型、网络的社会学模型、地震的地质学模型、生态位形成的生态学模型以及学习的心理学模型都体现了知识。这些模型能够解释和预测。化学键模型解释了为什么金属键会使我们无法将手伸进钢制的门，为什么当我们潜入湖水中时氢键会影响我们的体重。[8]

层次结构的基础就是智慧。智慧就是指识别和应用相关知识的能力。智慧需要多模型思维。有时，智慧体现在懂得如何选出最优模型，就好像将箭从箭袋中抽出来一样。还有时，智慧可以通过求出各种模型的平均结果来实现，这是在进行预测时的一种常见做法。采取行动时，有智慧的人都会应用多个模型，就像医生会让病人做好几种检查来帮助诊断一样。他们使用模型来排除某些行为、选择某些行为。有智慧的个

人和团队会有意让模型之间相互“对话”，探索不同模型之间的重叠和差异。

智慧包括选择正确的知识或模型。考虑一下这个物理问题：一个小小的毛绒玩具猎豹从一架飞在6千米高的飞机上掉下来，当它着地时会造成多大的伤害？学生可能已经掌握了引力模型和自由降落速度模型。这两个模型会给出不同的答案。引力模型的预测是，这个玩具猎豹会撕裂汽车的顶棚。自由降落速度模型的预测则是这个玩具猎豹的最高速度可以达到每小时16千米。[9]在这个问题上，智慧意味着，知道应该如何运用自由降落速度模型。事实上，站在地上的一个人，完全可以将这只柔软的毛绒玩具抓在手中。在此，不妨引用进化生物学家J.　B.　S.霍尔丹（J.　B.　S.　Haldane）的一段话来说明这个问题：“你可以将一只小鼠丢到一口深达千米的矿井，当它坠落到井底时，只要地面是相当柔软的，那么小鼠只会受到轻微的震荡，而且能够自行走开。但如果是大鼠的话就会摔死，人则会粉身碎骨，马更将尸骨无存。”

回到上面这个毛绒玩具的问题上来，要想得到正确的答案需要信息（这个玩具的重量）、知识（自由降落速度模型）和智慧（选择正确的模型）。商界和政界领袖也依靠信息和知识做出明智的选择。例如，2008年10月9日，冰岛的货币冰岛克朗（króna）开始自由落体般的急剧贬值。当时的软件巨头甲骨文公司（Oracle）的财务主管埃里克·鲍尔（Eric　Ball）必须做出一个决定。就在几个星期之前，他刚刚处理了国内住房抵押贷款危机带来的冲击。冰岛的情况引发了国际关注，而甲骨文公司持有数十亿美元的海外资产。鲍尔先考虑了关于金融崩溃的网络传染模型，然后他又考虑了讨论供给和需求的经济学模型（在这种模型中，价格变化的幅度与市场冲击的大小相关）。2008年，冰岛的国内生产总值仅为120亿美元，只相当于麦当劳公司6个月的销售收入。事后，鲍尔回忆当时的思考过程：“冰岛的经济规模比美国弗雷斯诺市还要小呢。回去工作吧，不用多管。”[10]

要理解这个例子，或者理解多模型思维方法，关键是要认识到鲍尔并没有去探索过多的模型，他找到了一个模型来支持已经决定采取的行动。是的，他没有尝试很多模型后找到一个能证明自己行为合理性的模型。相反，他只评估了两个可能有用的模型，然后选择了一个更好的模型。鲍尔拥有正确的信息（冰岛很小），选择了正确的模型（供需模型），并做出了一个明智的选择。

接下来，我们重新反思两个历史事件来说明如何让多个模型展开“对

话”。这两个历史事件是：2008年的全球金融市场崩溃，它使总财富（或者说至少是人们所认定的总财富）减少了数万亿美元，进而导致了长达4年之久的全球经济衰退；以及1961年的古巴导弹危机，它几乎引发了一场核战争。

对于2008年全球金融市场崩溃的原因，已经出现了多种解释：外国投资过多；投资银行过度杠杆化；抵押贷款审批过程缺乏监督；家庭消费者过分乐观的情绪；金融工具的复杂性；对风险的误解，以及贪婪的银行家明知泡沫存在却铤而走险并期望获得救助；等等。表面证据似乎与这些解释保持了一致：从外国流入了大量资金；贷款发起人发放了“有毒”（低质量）的抵押贷款；投资银行的杠杆率确实非常高；金融工具太过复杂导致大多数人无法理解；不少银行预计政府会出台救助计划；等等。通过模型，我们可以在这些解释之间加以“裁决”，可以分析其内在一致性：它们是否符合逻辑？我们还可以用数据进行校准、对推断进行检验。

经济学家罗闻全运用多模型思维方法，对关于这场危机的20种不同解释进行了评估。他发现，每一种解释都有不足之处。而且，没有理由认为投资者在明知自己的行为会导致全球危机时还会为泡沫作贡献。因此，泡沫的严重程度一定是出乎许多人的意料的。金融公司可能假定其他公司已经做好了尽职调查，而事实上并没有。回想起来，明显“有毒”的抵押贷款组合也找到了买家。如果全球金融市场崩溃成为定局，那么买家就不会存在。虽然杠杆率自2002年以来一直在上升，但却并没有比1998年的时候高出很多。而对于政府必定会救助银行的观点，雷曼兄弟银行的遭遇说明了一切：雷曼兄弟银行于2008年9月15日倒闭，它的资产超过6 000亿美元，这是美国历史上最大的破产案，然而政府并没有介入。

罗闻全认为，每种解释都包含了一个逻辑上的缺憾。从数据本身来说，没有任何一个解释是特别有根据的。正如罗闻全所总结的：“我们应该从一开始就努力对同一组客观事实给出尽可能多的解释，并寄希望于时间。当时机成熟的时候，关于这场危机更细致和更一致的解释就会浮现出来。”他还说：“唯有通过收集多样化且往往相互矛盾的解释，我们才能最终实现对危机更完整的理解。”任何单个的模型都是不足的。[11]

在《决策的本质》（*Essence of Decision*）一书中，美国政治学家格雷厄姆·艾利森（Graham Allison）采用多模型思维方法解释了古巴导弹危机。1961年4月17日，一支由美国中央情报局训练出来的半正规武装

队伍在古巴海岸登陆，企图推翻菲德尔·卡斯特罗的政权，加剧了美国与古巴的盟友苏联之间的紧张关系。作为回应，时任苏联总理尼基塔·赫鲁晓夫将短程核导弹运到了古巴。而时任美国总统约翰·肯尼迪则以对古巴的封锁作为回应。最终，苏联做出让步，危机结束了。

艾利森用三个模型解释了这个事件。首先，他运用理性行为者模型（rational-actor model）阐明，肯尼迪当时有三种可能的行动：发动核战争、入侵古巴或者进行封锁，最终他选择了封锁。理性行为者模型假设肯尼迪为每种行动绘制了一棵博弈树，并附上苏联可能做出的反应，然后，肯尼迪根据苏联的最优反应来思考自己的行动。例如，如果肯尼迪选择发动核战争，那么苏联就会反击，最终可能会造成数百万人死亡。如果肯尼迪决定封锁古巴，他就会使古巴人挨饿，而苏联则可能选择撤退或发射导弹。考虑到这个选择，苏联应该让步。这个模型揭示了核心策略逻辑（central strategic logic），并为肯尼迪大胆选择封锁古巴提供了合理的理由。

然而，尽管如此，像所有模型一样，这个模型也是错误的。它忽略了一些重要的相关细节，使它乍看起来比实际情况更好。这个模型也忽略了苏联已经将导弹运入古巴这个事实。如果苏联是理性的，他们应该会和肯尼迪一样画出博弈树，并认识到他们必须拆除导弹。理性行为者模型也无法解释为什么苏联没有将导弹藏起来。

其次，艾利森用组织过程模型（organizational process model）解释了这些不一致性。缺乏组织能力是苏联未能隐藏导弹的原因。这个模型也可以解释为什么肯尼迪选择封锁古巴，因为当时美国空军不具备在一次打击中就摧毁导弹的能力。即便只剩下一枚导弹，也会造成数百万美国人的伤亡。艾利森巧妙地结合了这两个模型。来自组织过程模型的洞察力，改变了理性选择模型（rational-choie model）中的结果。

最后，艾利森又使用了政府过程模型（governmental process model）。之前的两个模型都将国家化约为它们的领导者：肯尼迪代表美国行动，赫鲁晓夫代表苏联行动。政府过程模型则认为，肯尼迪不得不与国会抗衡，而赫鲁晓夫则必须维持支持自己的政治基础。因此，赫鲁晓夫在古巴部署导弹是一种力量的宣示。

艾利森这本书分别展示了模型本身以及模型之间对话的威力，每一个模型都能使思路变得更加清晰。理性行为者模型确定了导弹到达古巴后可能采取的行动，并帮助我们看清了这些行动的含义。组织过程模型让我们注意到了是组织而不是个人在实施这些行动。政府过程模型则突

出了入侵的政治成本。在通过所有这三个视角评估了这个事件后，我们就有了更全面、更深刻的理解。所有模型都是错的，但是同时运用多个模型确实非常有用。

在这两个例子中，不同的模型解释了不同的因果因素。此外，多模型思维方法也可以专注在不同的尺度上。在一个经常被人提及的故事中，一个孩子声称地球是驮在一头巨大的大象背上的。一位科学家问这个孩子，那么大象又是站在什么东西上呢。孩子回答道："一只巨大的乌龟的背上。"然后，科学家继续问，孩子继续答。不难预料接下来会发生的事情，孩子的回答是："你不要再问啦！乌龟驮乌龟，一直驮下去！"[12]

如果我们这个世界真的就是通过乌龟驮乌龟这样维持着的，或者说，如果这个世界是自相似（self-similar）的，那么最顶层的模型将适用每个层面。但是经济、政治世界和社会都不可能是这样的乌龟队列，大脑也不可能。在亚微米水平上，大脑由构成突触的分子组成，突触组成了神经元，神经元在神经元网络中结合。不同的神经元网络相互重叠，具体模式可以通过脑成像技术来加以研究。这些神经元网络存在的层级低于功能性系统（如小脑）。既然大脑在每个层级都有所不同，我们就需要多个模型，而且这些模型也各不相同。表征神经元网络稳健性的模型与用于解释脑细胞功能的分子生物学模型几乎没有任何相似之处，而后者又与用于解释认知偏差的心理学模型有所不同。

多模型思维的成功取决于一定程度的可分离性。在分析2008年金融危机的成因时，我们需要依赖外国人购买资产模型、资产组合模型、金融杠杆模型等多个模型。艾利森在根据博弈论模型进行推导时，不需要考虑组织过程模型。与此类似，在研究人体时，医生会将骨骼系统、肌肉系统、大脑系统和神经系统分开。也就是说，多模型思维并不要求这些不同的模型将系统分割为互不相关的部分。面对一个复杂的系统，用柏拉图的话来说，我们不能"将整个世界雕刻在关节上"。但是，我们可以部分地将主要的因果关系分离出来，然后探讨它们是如何交织在一起的。在这个过程中，我们将发现经济、政治和社会系统产生的数据会表现出一致性。这样一来，社会数据就不会再像家里养的猫一样吐出令人费解的毛球序列了。

做一个多模型思考者

现在总结一下。我们生活在一个充斥着信息和数据的时代。同时，这些数据得以产生的技术条件还极大地缩短了时间和空间上的距离。它们让经济、政治和社会行动者变得更加敏捷，能够在一瞬间就对经济和政治事件做出反应。它们还增加了连通性，因而也增加了复杂性。我们面临着一个由技术引发的悖论：在我们对世界的了解变得更多、更深入的同时，这个世界也变得更加复杂了。考虑到这种复杂性，任何单个模型都更有可能遭到失败。当然，我们不应该抛弃模型，恰恰相反，我们应该将逻辑一致性置于比直觉更优先的位置；我们不能满足于双重模型、三重模型甚至四重模型，我们要成为多模型思考者。

要成为一个多模型思考者，必须学习掌握多种模型，我们可以从中获得实用的知识，需要理解对模型的形式化描述，并知道如何应用它们。当然，我们也不一定非要成为专家不可。因此，这本书在可阅读性和论证深度之间做了一些权衡，它既可以作为学习资源也可以作为学习指导，书中对各个模型的正式描述都放在独立的专栏中。我还保证不会出现一行接一行都是方程式的情况，如果那样的话，即便是最专注的读者可能也无法忍受。不过，本书还是包括了少数几处包含方程式的论述，但它们都是容易理解的，也是应该被掌握的。构建模型是一门艺术，只能通过不断实践才能熟练掌握，这不是一项以观赏为目的的活动，需要刻意地练习。在建模中，数学和逻辑扮演着专家教练的角色，它们会纠正我们的缺漏。

本书其余各章安排如下。第2章和第3章讨论了多模型思维方法，第4章讨论了对人类建模的挑战。接下来的20几章，每章分别讨论一个模型或一类模型。由于一次只讲解一个模型，所以可以非常方便地将模型的假设、含义和应用厘清。这种章节结构也意味着，我们既可以阅读纸质书，也可以阅读电子书，而且可以直接去阅读与自己感兴趣的模型相关的章节每一章，我们都会应用多模型思维方法去解决各种各样的问题。本书最后给出了两个深度分析：一是针对类药物流行的现象，另外一个则涉及收入不平等问题。

02 模型的7大用途

了解现实就意味着构建转换系统，这些转换系统或多或少都必须与现实相对应。

让·皮亚杰（Jean Piaget）

在本章中，我们定义了模型的类型。人们通常认为，模型就是对世界的简化。是的，模型可以是对世界的简化，但是模型也可以采用类比的形式，或者，模型本身可能就是为探索思想和总结观点而构建的虚拟世界。在本章中，我们还描述了模型的7大用途。在学校里，我们应用模型来解释数据。在现实世界中，我们应用模型来预测、设计和采取行动，也可以使用模型来探索新思想和新的可能性，还可以利用模型来交流思想、增进理解。

模型的价值还体现在，它们能够把特定结果所需要的条件清晰地揭示出来。我们所知道的大多数结论都只是在某些情况下成立。例如，三角形最长边的平方等于另两边平方之和这个结论，只有当最长边是直角的对边时才成立。模型还可以揭示直觉结论可能成立的条件。我们可以分析传染病在什么情况下会传播、市场在什么条件下能正常运行、投票在什么环境下能够得到好的结果、群体在什么条件下能够给出准确预测……。这些都不是确定的事件。

本章分为两部分。在第一部分，我们描述了构建模型的3种方法。在第二部分，我们介绍了模型的7大用途：推理（reason）、解释（explain）、设计（design）、沟通（communicate）、行动（act）、预测（predict）和探索（explore）。这些用途的首字母，构成了一个缩略词“REDCAPE”。这个缩略词的字面含义为“红色披风”，提醒我们：多模型思维可以赋予我们强大的力量。[1]

构建模型的3种方法

要构建一个模型，我们可以在如下所述的3种方法中选择一种。

构建模型的第一种方法是具身法（embodiment approach）。用这种方法构建的模型包括重要部分，同时对于不必要的维度和属性，要么

剥离，要么将它们整合在一起考虑。生态沼泽模型、关于立法机构和交通系统的模型都是用这种方法构建的，气候模型和大脑模型也是如此。

构建模型的第二种方法是类比法（analogy approach），可以对现实进行类比与抽象。我们可以将犯罪行为传播类比为传染病传播，将政治立场的选择类比为在一个左-右连续线段上的选择。球形牛是类比方法的一个最直观的例子：为了估计一头牛身上牛皮的面积，我们会假设那头牛的形状是球形的。之所以要这样做，是因为微积分教科书所附积分表中的公式，会出现tan(x)和cos(x)，但是不会出现类似cow(x)这样的东西。[2]

相比而言，具身法更强调现实主义，而类比法则致力于刻画过程、系统或现象的本质。当一位物理学家假设不存在摩擦，同时又以其他方式做出符合现实的假设时，他所采用的就是体现法。当一位经济学家将相互竞争的公司视为不同的物种并在此基础上定义产品利基时，就是在做类比，用一个模型来表示不同的系统。但是，在具身法与类比法之间并没有一条明确的界限。例如，关于学习的心理学模型，在给不同的备选项分配权重时，往往会合并考虑多巴胺反应与其他因素，这种模型还会用我们在不同备选项之间进行权衡的方案做类比。

构建模型的第三种方法是另类现实法（alternative reality approach），也就是有意不去表征、不去刻画现实。这类模型可以作为分析和计算的“演练场”，我们可以利用这类模型探索各种各样的可能性。这种方法使我们能够发现适用于物理世界和社会世界之外的一般结论。这类模型有助于我们更好地理解现实世界中各种约束条件的含义，比如如果能够通过空气安全有效地传输能量，那么将会怎样？这类模型还允许我们进行现实世界中不可能的（思想）实验：如果我们能够加快大脑的进化，那么将会怎样？本书包含了不少这种类型的模型，其中一个是“生命游戏”（Game of Life），它是一个很大的棋盘，棋盘上的每一个方块要么是活的（黑色），要么是死的（白色），并根据某个特定规则在生死之间切换。虽然这个模型与现实世界并不一致，但是它能够帮助我们加深对自组织、复杂性现象的认识，甚至是许多关于生命本身的洞见。

无论是表征更复杂的现实世界、创造一个类比，还是建立一个用来探索思想的虚拟世界，任何一个模型都必须是易于处理且便于交流的。我们能够用形式化的语言对模型编码，比如数学符号或计算机代码。在描述模型时，我们不能在不给出正式描述的情况下直接抛出诸如信念或

偏好之类的东西。信念通常可以表示为一系列事件或先验的概率分布。而偏好则可以用多种方式来表示，比如用对一组备选项的排序或者一个数学函数来表示。

易于处理则是指适合分析的性质。在以往，分析依赖于数学运算或逻辑推理，因此建模者必须能够证明论证中的每一个步骤。这个约束条件导致了一种崇尚极致简约模型的“审美倾向”。神学家、哲学家奥卡姆的威廉（William of Ockham）提出了流传至今的“奥卡姆剃刀”原则：如无必要，勿增实体（Plurality must never be posited without necessity）。爱因斯坦则把“奥卡姆剃刀”原则进一步阐释为：事情应该力求尽可能简单，但是不可过于简单化。不过到了今天，当遇到用解析方法难以处理的问题时，我们还可以求助计算方法，可以构建由许多不断变化的组件的精细模型，而无须考虑解析上是否易于处理。科学家在构建全球气候模型、大脑模型、森林火灾模型和交通模型时，就采用了这种方法。当然，他们仍然不会忘记“奥卡姆剃刀”原则，只不过已经认识到“尽可能简单”还会要求很多不断变化的组成部分。

模型的7大用途

模型有几十种用途，不过在这里，我们只专注讨论其中的7种用途：推理、解释、设计、沟通、行动、预测和探索。

模型的7大用途（**REDCAPE**）

推理：识别条件并推断逻辑含义。

解释：为经验现象提供（可检验的）解释。

设计：选择制度、政策和规则的特征。

沟通：将知识与理解联系起来。

行动：指导政策选择和战略行动。

预测：对未来和未知现象进行数值和分类预测。

探索：分析探索可能性和假说。

REDCAPE：推理

在构建模型时，我们要先确定最重要的行为人（行动者）、实体以

及相关特征。然后，描述这些组成部分如何互动和聚合，我们能够推导出一些东西，并说明原因何在。这样一来，也就提高了我们的推理能力。虽然，能够推导出的东西取决于我们的假设，但是我们通过模型发现的绝不仅仅是重言式（tautology）[(2)]。因为我们很少能仅凭检验推断出假设的全部影响，我们需要形式逻辑。逻辑还可以揭示不可能性和可能性。利用模型进行推理，我们可以得到精确的，甚至是令人出乎意料的关系。我们可以发现自身直觉的制约性。

阿罗定理（Arrow's Theorem）就是一个可以说明逻辑如何揭示不可能性的极佳例子。这个模型解决了个人偏好是否集结为集体偏好的问题。在这个模型中，偏好表示为各备选项之间的排序。以对餐馆进行排名为例，假设有5家意大利餐馆，分别用字母A到E表示，这个模型允许120种排序中的任何一种。阿罗要求集体排序是单调的（如果每个人都将A排在B之前，那么集体排序也是如此）、独立于无关的备选项（在其他备选项的排名发生了变化的情况下，如果任何人对A和B的相对排名都没有发生变化，那么A和B在集体排名中的顺序也不会改变），且是非独裁的（没有任何一个人能够决定集体排序）。然后阿罗证明，如果允许任何偏好都存在，那么就不存在集体排序。[3]

逻辑也可以揭示悖论。利用模型，我们可以证明，每个亚种群中的女性人口比例大于男性，但是在整个种群中却是男性人口的比例更高，这种现象被称为“辛普森悖论”（Simpson's paradox）。在现实世界中，这种情况已经发生过了：1973年，加州大学伯克利分校的绝大多数院系都录取了更多的女生，但是从总体上看，它却录取了更多的男生。模型还表明，两个没有胜算的赌局，当交替轮流进行时，是有可能带来正的预期回报的，这就是人们熟知的“帕隆多悖论”（Parrondo's paradox）。通过模型，我们可以证明，在向网络中添加节点的同时，是可以减少连接所有节点所需边的总边长的。[4]

需要注意的是，我们不能把上面这些模型的例子简单地视为数学上的新奇事物。事实上，每一个模型都有很大的实际应用价值：提高女性在人口中比例的努力可能会适得其反；将没有机会赢利的投资适当地组合起来可能会带来收益；电线、管道网、以太网线路或道路网的总长度可以通过增加更多的节点来减少等。

逻辑也可以揭示数学关系。根据欧几里得定理，三角形可以由任意两个角和一条边，或任意两条边和一个角唯一确定。根据对消费者和公司行为的标准假设，当市场上有大量的相互竞争的企业时，价格等于边

际成本。但是，这里也会出现一些出乎意料的结果，其中一个是所谓的“友谊悖论”（friendship paradox），它说的是，在任何一个由友人组成网络中，平均而言，一个人的朋友拥有的朋友要比这个人更多。

“友谊悖论”之所以会出现，是因为非常受欢迎的那些人有更多的朋友。图2-1显示的是扎卡里（Zachary）的空手道网络。在图中，黑色的圆圈所代表的人有6个朋友，这些朋友用灰色圆圈表示，他的朋友们平均每个人有9个朋友。

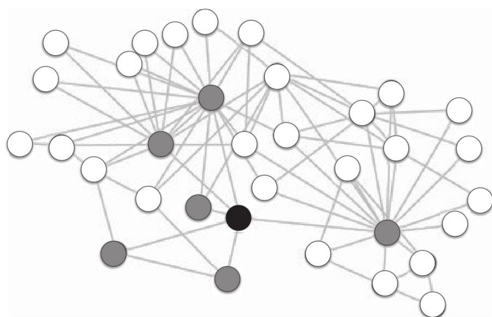


图**2-1**　友谊悖论：朋友拥有的朋友比自己多

在整个网络中，34人中有29人拥有比他们自己更受欢迎的朋友。[5]稍后在下文中，我们还将了解到，只要加入更多的假设，那么大多数人的朋友平均来说会比自己更加好看、更加善良、更加富有、更加聪明。

最重要的是，逻辑还揭示了真理的条件性。政客可能会声称降低所得税会通过促进经济增长，从而增加政府收入。但是，根据政府收入等于收入水平乘以税率的基本模型，我们很容易就可以证明，只有当收入的百分比增幅超过了减税的百分比时，政府收入才会增加。[6]因此，收入税减少10%的政策，只有在它能够导致收入增幅超过了10%时，才能带来政府收入的增加。政客的逻辑只适用于某些特定条件，而模型就能将这些条件识别出来。

当我们将模型中推导出来的主张与叙述性主张进行比较时，这种“条件性”的威力将会变得更加明显，即便后者有经验证据支持时也是如此。我们先来考虑一下这个管理名言：重要的事情先做（first thing first）。它说的是，在面对多项任务时，你应该首先完成最重要的那项任务。这个原则有时也被称为“大石头优先”原则，意思是当你要将一些大小不一的石头装入一只桶中时，你应该先装入大石头，如果你先放入小石头，那么大石头就放不下了。

“大石头优先”原则，是专家从观察中总结出来的，在许多时候确实不失为一个很不错的原则，但是它也不是无条件的。基于模型的方法将会先对任务提出具体的假设，然后推导出最优规则。例如，在“装箱问题”（bin packing problem）中，必须将一系列不同大小或不同重量的物体装入容量有限的箱子中，目标是保证所用的箱子尽可能少。

不妨想象这样一个场景：你准备搬家，要把家中的所有东西打好包，放入若干个50厘米×50厘米的箱子里。把你的所有东西按大小排好序，然后将每一件东西放入第一个有足够空间的箱子，这种方法称为“首次适应算法”（first fit algorithm），事实证明相当有效。这就是说，“大石头优先”原则的效果非常不错。

但是，假设我们要考虑一个更加复杂的任务：在国际空间站上，为若干研究项目分配空间。每个项目都对有效载荷重量、空间大小和动力有一定要求，对宇航员的时间和认知能力也有自己的要求。而且，每个项目都有做出科学贡献的潜在能力。在这个问题中，即便我们想出了一个衡量这种“大石头”（重要性）的方法，对上面这些属性求加权平均值的权重，但在给定的相互依赖性的维度下，“大石头优先”原则也已经被证明是一个相当糟糕的原则。更复杂的算法以及可能的市场机制则会更好地发挥作用。[7]因此，在某些条件下，“大石头优先”原则可能是一个很好的原则。但是，在另外一些条件下，“大石头优先”原则就不行了。通过利用模型，我们可以划出一条界线：什么时候应该采用、什么时候不能采用。

形式主义的批评者声称，说到底，模型只不过是对我们已经知道的东西进行了重新包装而已，只不过是将旧酒倒入闪闪发亮的“数学”新瓶中而已。这些批评者可能会说，难道我们不知道“三个臭皮匠，赛过诸葛亮”吗？难道我们不知道“三思而不行，终将无所得”吗？我们不需要模型就能知道这些道理。他们还认为，我们可以通过阅读《荷马史诗》中奥德赛将自己绑在桅杆上的故事，懂得承诺的价值。

但这些批评者没有认识到，从模型中得出的推论总是采用条件判断形式：如果条件A成立，那么可以得出结果B。例如，如果你要装箱，而大小是唯一的约束条件，那么就先装好最大的东西。我们从经典文献和伟大思想家的名言中吸取的教训却通常不包括任何条件。如果我们试图依据这种“原则”来生活或管理他人，就肯定会迷失在众多意思相反的谚语海洋当中，既然有“三个臭皮匠，赛过诸葛亮”，也会有“厨子多了烧坏汤”（表2-1）！

表2-1　相互对立的谚语

谚语	相反的谚语
三个臭皮匠，赛过诸葛亮 （Two heads are better than one）	厨子多了烧坏汤 （Too many cooks spoil the broth）
三思而不行，终将无所得 （He who hesitates is lost）	一针及时顶九针 A stitch in time saves nine
破釜沉舟 （Tie yourself to the mast）	留条后路 （Keep your options open）
“完美”是“优秀”之敌 （The perfect is the enemy of the good）	要做就要尽善尽美，不然不如不做 （Do it well or not at all）
事实胜于雄辩 （Actions speak louder than words）	笔尖强过干戈 （The pen is mightier than the sword）

而在模型中，我们可以在给定的假设下证明定理。相反的谚语经常共存，但是相反的定理却不会出现。两个定理，如果对何为最优行动有不同看法，必定会做出不同的预测；或者，给出了不同解释的定理必定有不同的假设。

REDCAPE：解释

模型为经验现象提供了清晰的逻辑解释。经济学模型解释的是价格变动和市场份额等现象；物理学模型可以解释坠落物体的轨迹和轨迹形状的变化；生物学模型可以解释物种的分布；流行病学模型解释了传染病传播的速度和模式；地球物理学模型能够解释地震的大小和分布。

模型可以解释点值（point values）和点值的变化。例如，某个模型可以解释五花肉期货的当前价格以及过去6个月来价格上涨的原因。另一个模型可以解释为什么美国总统会任命持温和立场的最高法院法官，以及为什么美国总统候选人会向左翼或右翼靠拢。模型还可以解释形状：关于思想、技术和传染病传播的模型，都会产生S形的采用曲线（或传染曲线）。

我们在物理学中学到过不少模型，例如玻意耳定律，这个定律告诉我们，氧气的压力乘以体积等于一个常数（$pV=k$），这个定律非常完美地解释了许多现象。[8]如果知道了体积，就可以估计出常数k，然后就可以解释压力p，或者预测作为V和k的函数的压力p。这个模型的准确性可以归因于如下事实：气体由大量存在的简单成分组成，而且遵循一个固定不变规则，即任何两个氧分子在相同情况下必定遵循相同的物理定

律。氧分子的数量如此之多，以至于统计上的平均值可以抹去任何随机性。

但是，大多数社会现象都不具备这三种性质：社会行动者是异质性的、互动是在小群体内展开的、行为人也不遵守固定的规则。此外，人还会思考。更加重要的是，人会对社会上的风吹草动做出反应，而这就意味着行为变化可能是无法相互抵消的。因此，社会现象要比物理现象更加难以预测。[9]

最有效的模型既能解释简单的现象，也能解决令人费解的问题。教科书中关于市场的经典模型能够解释为什么对于像鞋子或薯片这样正常商品需求的意外增加，会在短期内提高它们的价格，这是一个非常直观的结果。这些模型还可以解释，为什么从长期来看，需求增加对价格的影响会小于生产商品的边际成本的影响。需求的增加甚至有可能会导致价格下降，这种现象在规模收益增加的情况下确实会出现。这无疑是一个更令人惊讶的结果。这些模型还可以解释一些悖论，例如水和钻石悖论：钻石只具有很小的实用价值，但是价格却很高；水虽然是人类生存的必需品，但价格却很低。

有人说，模型可以解释任何东西。这种说法没有错，模型确实可以。然而，基于模型的解释必须包括正式的假设和明确的因果链条，而且这些假设和因果链条都要面对数据。例如，有个模型说，用低被捕概率可以解释犯罪率的居高不下，这样的模型就是可检验的。

REDCAPE：设计

模型还可以通过提供框架来帮助设计，因为只有在适当的框架内我们才可以考虑不同选择的含义。工程师使用模型设计供应链；计算机科学家使用模型设计Web协议；社会科学家使用模型设计制度。

1993年7月，一群经济学家在位于加利福尼亚州帕萨迪纳市的加州理工学院开会，设计一种拍卖方法，拍卖对象是手机所用的电子频谱。在那之前，美国政府一直将频谱的使用权分配给大型公司使用。1993年通过的《统一综合预算协调法案》（*Consolidated Omnibus Budget Reconciliation Act*）则允许政府拍卖频谱以筹集资金。

从一座信号塔发射的无线电信号只能覆盖一定的地理区域。因此，政府可以出售各个特定地区的许可证，例如，俄克拉何马州西部、加利福尼亚州北部、马萨诸塞州、得克萨斯州东部等。这就提出了一个设计

问题。一家公司所拥有的任何一张给定的许可证的价值，取决于该公司得到的其他许可证。例如，加利福尼亚州南部许可证对于拥有加利福尼亚州北部许可证的公司来说更有价值。经济学家将价值的这种相互依赖性称为外部性。这里的外部性有两个主要来源：建设成本和广告市场。持有相邻地区的许可证意味着更低的建设成本和利用重叠的媒体市场的潜力。

这种外部性对同时举行的拍卖提出了挑战。一家试图赢得一组许可证的公司可能会在其中某一张许可证的拍卖中输给另一个竞标人，并因此而失去所有外部性，也就是可以带来的收益。那样的话，这家公司就可能会希望退出其他许可证的拍卖。然而另一方面，连续拍卖也有一个缺点。竞标人在前面的许可证拍卖中会出低价，以对冲在后面的拍卖中竞买失败可能导致的损失。

成功的拍卖制度设计必须符合这样一些要求：不会受策略性操纵的影响、能够产生有效率的结果，同时又容易被拍卖参与者所理解。为此，参加加州理工学院会议的那些经济学家，利用博弈论模型分析了策略性竞标人可能会利用的各种特征，采用计算机模拟比较了各种设计方案的效率，还通过统计模型选择了真人实验的参数。最终，他们设计出了一种多轮拍卖方法，做到了允许参与者退出竞标并禁止早期竞标人掩盖真实意图。事实证明，这是成功的。过去的30年以来，美国联邦通信委员会已经使用这种拍卖方法筹集了将近600亿美元资金。[10]

REDCAPE：交流

由于创造了一种共同的表示方法，模型能够有效地改进交流。模型要求对相关特征及其关系给出正式的定义，这使我们能够精确地进行交流。例如，模型$F=ma$，涉及3个可测量的量——力、质量和加速度，并将它们之间的关系用方程式的形式表示出来。每一项都可以表示为可测量的单位，因而可以很方便地就这个模型进行交流，而不必担心会有什么误解。相比之下，“更大、更快的东西会产生更大的力”这种说法的准确度却要低得多。因为这需要翻译，而翻译会令很多人“迷失方向”。“更大”指的是重量还是体积？“更快”指的是速度还是加速度？“力”指的是能量还是力？“更大”和“更快”的结合又怎么产生“力”呢？对这种说法的定义，也有不同的方向：可以将“力”写为重量与速度之和（$P=W+V$）、重量与速度之积（$P=W \quad V$），又或者写为重量与加速度之和（$P=W+A$）……

当我们根据可复制性的要求，给像“政治意识形态”这样的抽象概念下了一个定义之后，这些概念也就具有了与质量和加速度等物理概念等量齐观的某些特征。我们可以通过一个模型给出这样的论断，根据他们的投票记录，某个政客比另一个政客更“自由”（“保守”）。然后我们可以准确无误地用这种论断与他人交流。“自由”是有明确定义的，而且是可度量的。其他人可以使用相同的方法去对其他政客进行比较。当然，投票记录可能不是衡量“自由”与“保守”的唯一标准。这时我们可以构建出第二个模型，根据演讲的文本分析来分配意识形态立场。有了这种模型，也可以将我们所说的更加“自由”的意思准确无误地传达给其他人。

很多人都低估了交流对人类社会进步的影响。一个无法交流的思想，就像一棵淹没在森林中的树，没有人会注意到它。启蒙时代显著的经济增长在很大程度上取决于知识的可传播性（知识通常表现为模型形式）。事实上，有充分证据表明，在那个时代，思想的可传播性对经济增长的贡献，比教育水平还要大。其中一个有力的证据是，在18世纪的法国，各城市的经济增长与狄德罗（Diderot）的《百科全书》（*Encyclopédie*）的订阅数量之间的相关性，远远高于与识字率之间的相关性。[11]

REDCAPE：行动

弗朗西斯·培根曾经这样写道：“人生的伟大目标，不在于知，而在于行。”良好的行动需要良好的模型。政府、企业和非营利组织都要使用模型来指导行动。无论是提高价格（降低价格）、开设新的分支机构、兼并其他公司、提供全民医疗保健，还是资助某个课外计划，决策者都要依赖模型。在最重要的行动中，决策者要使用多个复杂的模型，模型与数据紧密相关。

2008年，作为《问题资产救助计划》的一部分，美国联邦储备银行提供了1　820亿美元的金融救助款，以拯救跨国保险公司美国国际集团（AIG）。根据美国财政部的报告，政府之所以决定拯救美国国际集团，是“因为它在金融危机期间如果破产，就会对我们的金融体系和经济产生破坏性影响”。[12]救助的目的不是为了拯救美国国际集团本身，而是为了支持整个金融体系。每天都有企业破产，但是政府通常不会介入。[13]

根据《问题资产求助计划》做出的每一项具体决策都是以特定模型为基础的。图2-2显示了国际货币基金组织给出的一个网络模型。在这

里，节点（圆圈）代表金融机构，边代表这些金融机构的持有资产价值之间的相互关系。连接的颜色和宽度代表相关性的强度，更深和更粗的线条意味着更大的相关性。[14]

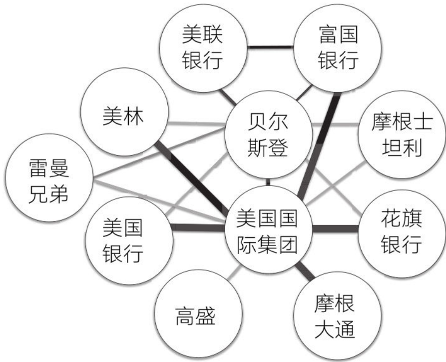


图2-2　金融机构的网络模型

从图2-2可见，美国国际集团在这个金融机构网络中占据了中心位置，因为它向其他公司出售保险。如果这些公司的资产价值蒙受了损失，美国国际集团根据承诺要向它们支付赔偿金。这就是说，如果资产价格下跌，那么美国国际集团就欠了这些公司钱。这个网络的隐含义是，如果美国国际集团破产了，那么与它相关的公司也会破产，从而很可能会引发一连串的破产。通过稳定美国国际集团，美国政府可以为网络中的其他公司的市场价值提供支持。[15]

图2-2也有助于解释为什么政府会让雷曼兄弟公司倒闭：因为雷曼兄弟公司并没有在网络中占据中心地位。我们不能让历史重演，所以我们无法确知美国联邦储备委员会当时是不是采取了正确的行动。但是我们确实知道，雷曼兄弟公司的破产没有导致金融体系崩溃，而且政府向美国国际集团的贷款还为它带来了230亿美元的利润。

指导行动的模型通常依赖于数据，但并不是全部模型都依赖于数据。大多数政策模型都需要使用数学公式，但也并非总是如此。过去，决策者也曾经建立过物理模型。在20世纪中期，菲利普斯（Phillips）为英国经济构建了一个水力模型，它一度被用于考虑政策选择。另外，关于旧金山湾的物理模型，也对终结将旧金山湾改造为淡水湖的计划起到了很大作用。[16]密西西比河流域模型水道试验站（Mississippi River Basin Model Waterways Experiment Station）建造的流域缩微模型，位于密西西比州克林顿市附近，占地近80万平方米，按1：100的比例完整复制了整个流域。这个模型可以检验建造新水坝和水库对流域上游和下

游地区的效应，因为放出来的水会遵循物理结构中的物理定律。这样的物理模型中，嵌入数据的物理实体和物理定律自然会“完成”逻辑推理。

到目前为止，我们举的例子都是关于组织如何使用模型来采取行动的，个人当然也可以这么做。在日常生活中，当我们准备采取某个重要行动时，也应该使用模型。例如，在决定购买房屋、更换工作、回到大学攻读更高的学位，或者在决定是购买还是租赁汽车时，都可以使用模型来指导决策。用到的模型可能只是一些定性的模型而不一定有相应的数据支持，但是多模型思维会“迫使”我们向自己提出一些重要的问题。

REDCAPE：预测

模型长期以来被用来预测。天气预报员、专家、顾问和许多国家中央银行行长，都在使用模型进行预测。警察机构和情报部门也使用模型预测犯罪行为，流行病学家则使用模型预测下个季节哪种流感病毒将最为流行。现在，随着数据可得性的提高和精细度的改进，利用模型进行预测的做法变得更加常见了。例如，Twitter上的跟帖和谷歌上的搜索关键词，都已经被用于预测消费趋势和潜在的社会活动了。

模型既可以用来预测特定的个别事件，也可以用来预测一般趋势。2009年6月1日，法国航空公司的AF 477航班在从里约热内卢飞往巴黎的途中，在大西洋上空坠毁。在接下来的几天里，救援人员发现了一些漂浮的碎片，但是无法找到尸体。到7月份，飞机上的水下信标中的电池耗尽了电力，搜索不得不中止。一年后，伍兹霍尔海洋研究所（Woods Hole Oceanographic Institution）率领的搜寻队使用美国海军的侧扫声呐船和水下自动航行器进行了第二次搜索，也没有得到什么结果。于是，法国国家统计分析局（French Bureau d'Enquêtes et d'Analyses）不得不求助于模型。他们将概率模型应用于大海洋流，并识别出了一个坠毁的飞机最有可能沉没的矩形区域（面积并不很大）。根据模型给出的这个预测，搜索队在一个星期之内就找到了飞机残骸。[17]

过去，解释和预测往往是齐头并进的。解释电压模式的电气工程模型也可以预测电压大小，解释政客过去投票行为的空间模型也可以预测他们在未来的投票。运用原本用于解释的模型进行预测的最著名的一个例子是，法国数学家、天文学家奥本·勒维耶（Urbain Le Verrier）运用解释行星运动的牛顿定律，预测还存在另一颗行星，进而以此来解释天王星运行轨道的异常。勒维耶证明，那些轨道与太阳系外围地区存在另

一颗大行星时的轨道一致。1846年9月18日，勒维耶将预测发给了柏林天文台。5天后，天文学家就在勒维耶预测的那个位置上发现了海王星。

不过话说回来，预测毕竟是与解释不同的。有的模型可以用来预测，但是却不一定能解释什么。深度学习算法可以预测产品的销售情况、明天的天气变化、价格演变趋势和身体健康状况，但是它们几乎没有提供什么解释。这些模型类似于“嗅弹犬”。尽管这些狗可以利用它们灵敏的嗅觉系统确定一个包裹是不是包含着爆炸物，但是我们确实不应该要求它们解释为什么知道那里有炸弹，也不能去问它们工作原理是什么、怎样才能拆除炸弹。

此外还要注意到，有些模型有很强的解释力，但是在预测上却没有什么价值。板块构造论模型虽然可以解释地震是怎样发生的，但是却不能预测地震何时发生；动力系统模型虽然可以解释飓风是怎样形成的，但是却无法准确预测飓风什么时候袭来，也不能准确预测飓风的移动路径；生态模型虽然可以解释物种的形成的模式，但是却无法预测出现的新物种类型到底是什么。

REDCAPE：探索

最后，我们还会用模型来探索直觉。这种探索可能与政策相关：如果让所有城市公交车都免费，会怎么样？如果让学生自主选择作业来证实他们的课程成绩，会怎么样？如果在草坪上标出能量消耗数量，又会怎么样？我们可以提出很多假说，而且所有这些假说都可以用模型进行探索。我们还可以利用模型来探索某些在现实世界中不会出现的情况。如果法国生物学家拉马克（Lamarck）的观点是正确的、如果后天获得的性状真的可以遗传给我们的后代，那么那些把牙齿矫正好了的父母的孩子就再也不需要牙套了吗？在这样的世界还会发生什么？提出这样的问题并探索它们的含义可以帮助我们揭示进化过程的局限性。暂且将现实世界的约束丢到一边，可以极大地激发我们的创造力。也正是出于这个原因，批判性设计运动的许多倡导者都利用科幻小说来促进思考并提出了不少新的思想。[18]

探索有时还涉及对共同假设进行跨领域比较。例如，为了理解网络效应，建模者可能会从一系列程式化的网络结构入手，然后追问网络结构是不是会影响以及如何影响合作、传染病传播或社会动乱。又或者，建模者可能会将一系列学习模型应用于决策、双人博弈和多人博弈；但

是他们这样做的目的不是为了解释、预测、行动或设计，而只是为了探索和学习。

当我们在实践中应用一个模型时，也能以多种方式使用它。同一个模型既可以用来解释、预测，也可以用来指导行动。例如，2003年8月14日，俄亥俄州托莱多市附近，树木倒塌压断了电线，造成了局部电力中断，但是，由于监控软件出了故障，没有及时发出警报，让技术人员去对电力进行重新分配，最后导致一天之内，美国东北部和加拿大有超过5 000多万人遭受停电之苦。同一年，意大利和瑞士之间的电线受风暴袭击，导致6 000万欧洲民众无法用电。为此，工程师和科学家求助于将电网表示为网络的模型。这些模型不但有助于解释故障是如何发生的，而且有助于预测未来可能出现故障的区域。它们还能够识别出为了增强电网的稳定性，应该在哪些地方增加新的线路、新的变压器和新的电源，从而起到指导行动的作用。

将一个模型用于多种用途，正是本书中将会反复出现的一个主题。正如接下来将会看到的，一对多是运用多个模型来理解各种复杂现象这一中心主题的必要补充。

03　多模型思维

没有什么比现实主义更不真实了……细节令人困惑。只有通过选择、通过消除、通过强调，我们才能获得事物的真正意义。

乔治娅·奥·吉弗（Georgia O' Keeffe）

本章将通过科学的方法来引入多模型思维。我们先从孔多塞陪审团定理（Condorcet jury theorem）和多样性预测定理（diversity prediction theorem）入手讨论，这两个定理为证明多模型思维在帮助人们行动、预测和解释方面的价值提供了可量化的论据。需要指出的是，这两个定理可能夸大了许多模型的情况。为了说明原因，我们又引入了分类模型（categorization model），它将世界划分为一个个箱子。使用分类模型的目的是表明构建多模型可能会比预想的更难。然后，我们利用这类模型讨论了模型粒度（model granularity），也就是模型应该有多具体，并帮助我们决定是采用一个大模型还是多个小模型。选择取决于用途：在预测时，我们经常需要大模型；而在解释时，小模型则更好一些。

我们得到的结论解决了一个长期以来挥之不去的忧虑：多模型思维可能需要学习非常多的模型。是的，虽然我们必须学习掌握一些模型，但是并不需要学习像有些人想象的那么多。我们不需要掌握100个模型，甚至连50个也不需要，因为模型具有一对多的性质。我们可以通过重新分配名称、标识符，或者修改假设来将任何一个模型应用于多种情况。模型的这个性质很好地平衡了多模型思维的需求。事实上，在新的领域应用模型对创造力、开放性和怀疑精神的要求也非常高。我们必须认识到，并非每个模型都适合每项任务。如果一个模型无法解释、预测或帮助我们推理，那就必须将它放到一边，考虑其他模型。

这种一对多的技能，与许多人所认为的要成为一名优秀建模者所必需的数学和分析才能是不同的。一对多的过程对创造力的要求很高，它实际上相当于在问这样一个问题：对于随机游走，我能够想到多少种用途？作为这种创造力的一个例子，在本章的最后，我们将几何学中的面积公式和体积公式作为模型，解释了超级油轮的大小、评估了身体质量指数、预测了新陈代谢的比例……并解释为什么我们很少看到女性CEO。

孔多塞陪审团定理和多样性预测定理

现在来看看正式模型，它们有助于理解多模型思维的好处。在这些模型的情境下，我们描述了两个定理：孔多塞陪审团定理和多样性预测定理。

孔多塞陪审团定理是从一个解释多数规则长处的模型中推导出来的。在这个模型中，陪审员要做出要么有罪、要么无罪的二元决策。每个陪审员正确决策的时候比错误的时候多。为了将这个定理应用于模型集合而不是一组陪审员，我们将每个陪审员的决策解释为模型的一个类别。这种分类可以是行动（买入或卖出），也可以是预测（美国民主党胜出还是共和党胜出）。孔多塞陪审团定理告诉我们，通过构建多个模型并使用多数规则，将比只使用其中一个模型更加准确。这个模型依赖于世界状态（state of the world）的概念，它是对所有相关信息的完整描述。对于一个陪审团来说，世界状态包括了审判时呈现的所有证据。对于那些衡量某个慈善项目的社会捐献的模型来说，世界状态则可能与项目的团队、组织结构、运营计划以及项目所要解决的问题的特征或状况相对应。

孔多塞陪审团定理

总数为奇数的一组人（模型）将未知的世界状态分为真或假。每个人（模型）正确分类的概率为$p>1/2$，并且任何一个人（模型）分类正确的概率在统计上都独立于任何其他人（模型）分类的正确性。

孔多塞陪审团定理：多数投票正确的概率比任何人（模型）都更高；当人数（模型数）变得足够大时，多数投票的准确率将接近100%。

那么，如何将这个定理的原理应用于多模型方法呢？生态学家理查德·莱文斯（Richard Levins）对此给出了详细的阐述："因此，我们尝试用几个不同的模型来处理同一个问题，这些模型的简化方法各不相同，但都有一个共同的生物学假设。如果这些模型（尽管它们有不同的假设）都导致相似的结果，那我们就得到了一个强有力的定理，它基本上不受模型细节的影响。因此，我们的真理就是若干独立的谎言的交集。"[1]需要注意的是，在这里，莱文斯渴望达成一致的分类。当许多模

型都给出了相同的分类时，我们会信心大增。

多样性预测定理则适用于给出数值预测或估值的模型，它量化了模型的准确性和多样性对所有模型平均准确性的贡献。[2]

多样性预测定理

多模型误差＝平均模型误差-模型预测的多样性，即：

$$(\bar{M}-V)^2=\sum_{i=1}^{N}\frac{(M_i-V)^2}{N}-\sum_{i=1}^{N}\frac{(M_i-\bar{M})^2}{N}$$

在这里，M_i表示模型i的预测，$\bar{M}$等于模型的平均值，V等于真值。

多样性预测定理描述了一个数学恒等式。我们用不着费心检验，因为它总是成立。下面举一个例子来说明这一点。假设我们用两个模型来预测某一部电影会获得多少项奥斯卡奖。一个模型预测它将获得两项奥斯卡奖，另一个模型则预测它将获得8项。这两个模型预测的平均值，也就是多模型预测的结果等于5。如果最后这部电影获得了4项奥斯卡奖，那么第一个模型的误差等于4（2^2），第二个模型的误差等于16（4^2），而多模型误差则等于1，模型预测的多样性等于9（因为每个模型的预测与平均预测均相差3）。这样一来，多样性预测定理就可以表达为：1（多模型误差）＝10（平均模型误差）-9（模型预测的多样性）。

这个定理的原理在于，相反类型的误差（正负）会相互抵消。如果一个模型的预测值太高，同时另一个模型的预测值太低，那么这些模型就会表现出预测多样性。两个模型的误差相互抵消，模型的平均值将比任何一个模型更加准确。即便两个模型的预测值都太高，这些预测值的平均误差仍然不会比两个高预测值的平均误差更糟。

但是，多样性预测定理并不意味着任何不同模型的集合的预测必定是准确的。如果所有模型都有一个共同的偏差，那么它们的平均值也会包含那个偏差。不过，这个定理确实意味着，任何多样性的模型（或人）的集合将比其普通成员的预测更加准确，这种现象就是通常所说的“群体的智慧”（wisdom of crowds）。这是一个数学事实，它解释了计算机科学中集成方法（ensemble method）成功的原因，这种方法对多个分类加以平均，也解释了使用多个模型和框架进行思考的人比使用

单个模型的人预测的准确性更高的事实。任何一种看待世界的单一方式都会遗漏掉某些细节，使我们更容易产生盲点。单模型思考者不太可能准确预测到重大事件，例如2008年的金融危机。[3]

这两个定理为我们利用多个模型提供了令人信服的理由，至少在进行预测的情况下。然而，这个理由在一定意义上可能显得过强。孔多塞陪审团定理意味着，如果有足够多的模型，我们几乎永远不会犯错。多样性预测定理则意味着，如果能够构建一组多样的中等准确性的预测模型，我们就可以将多模型误差减少为接近于零。但是，正如接下来将会看到的，我们构建多个多样性模型的能力是有限的。

分类模型

为了说明为什么这两个定理可能会“夸大其词”，现在来讨论一下分类模型。这类模型为孔多塞陪审团定理提供了微观基础。分类模型将世界状态划分为不相交的。最早的分类模型可以追溯到古希腊时代。在《范畴篇》（*The Categories*）一书中，亚里士多德描述了对世界进行分类的10个范畴，包括了实体（substance）、数量（quantity）、地点（location）和状态（positioning）等，每个范畴都会创建不同的类别。

当我们使用一个普通名词时，“裤子”是一个类别，“狗”、“勺子”、“壁炉”和“暑假”也是如此。我们就是在使用类别去指导行动。我们按种族，比如意大利人、法国人、土耳其人或韩国人，来对餐馆进行分类，以便决定在哪里吃午餐；按照市盈率对股票进行分类，并根据市盈率高低买卖股票。当人们声称亚利桑那州的人口之所以增长是因为该州气候宜人时是在用分类方法进行解释。我们还使用类别进行预测，例如预计身为退伍军人的候选人在选举中会有更大的获胜机会。

我们还可以在智慧层次结构中解释分类模型的作用。对象构成了数据，将对象分为不同类别就能创造出信息，而将估值分配给各个类别则需要知识。为了评价孔多塞陪审团定理，我们依赖一个二元分类模型，它将对象或状态分为两个类别，一类标记为“有罪”，另一类标记为“无罪”。关键的思想是，相关属性的数量限制了不同类别的数量，因此也就限制了有用模型的数量。

分类模型

存在一组世界的对象或状态，每个对象或状态都由一组属性定义，每个属性都有一个值。根据对象的属性，分类模型M将对象或状态划分为一个有限的类别集$\{S_1, S_2, \ldots\ldots, S_n\}$，然后给每个类别赋值$\{M_1, M_2, \ldots\ldots, M_n\}$。

假设有100份学生贷款，其中有一半是按期还款的，另一半是违约的。我们知道每一笔贷款的两个信息：第一，贷款金额是否超过了5万美元；第二，贷款者主修的工科还是文科。这是两个属性。通过这两个属性，我们可以区分出4种类型的贷款：主修工科学生的大额贷款、主修工科学生的小额贷款、主修文科学生的大额贷款以及主修文科学生的小额贷款。

二元分类模型将上面这4种类型中的每一种都分为按期还款与违约。一种模型可能将小额贷款归为按期还款，将大额贷款则归为违约。另一种模型则可能将主修工科学生的贷款归为按期还款，将主修文科学生的贷款归为违约。我们有理由认为这两种模型中的任何一个都可能在超过一半的情况下是正确的，而且这两种模型大体上相互独立。

但是，当我们尝试构建更多的模型时就会出现问题。要将4个类别映射为两个结果，最多只有16个模型。上面这两个模型是其中的两个，它们将所有贷款分为按期还款或违约。剩下的14个模型中的每一个都有一个完全相反的模型，只要某个模型的分类是正确的，那么与之相反的那个模型的分类就是错误的。因此，在14个可能的模型中，最多只有7个可能在超过一半的情况下是正确的。而且，如果任何一个模型碰巧在一半的情况下是正确的，那么与它相反的模型也必定如此。

数据的维数限定了可以创建的模型数量，最多可以有7个模型。我们无法创建出11个独立的模型，更不用说77个了。而且，即使我们有更高维度的数据，比如，假设我们知道贷款者的年龄、平均成绩、收入、婚姻状况和住址，那么依赖这些属性的分类一定能产生准确的预测。每个属性子集都必须与贷款是否已经偿还相关，同时还必须与其他属性无关。这两者都是很强的假设。例如，如果收入、婚姻状况和住址是相互相关的，那么交换这些属性的模型也将是相互相关的。[4]在严格的概率模型中，独立性是合理的：不同的模型会产生独立的错误。运用分类模型的原理分析孔多塞陪审团定理的逻辑时，我们看到了构建多个独立模型的困难。

在试图构建一组多样性的、准确的模型时，也可能会遇到类似的困

难。假设我们想要构建一个分类模型来预测500个中型城市的失业率。一个准确的模型必须将这些城市划分为多个类别，以便让同一个类别中的城市具有相似的失业率，而且该模型必须能够准确地预测该类别的失业率。对于两个进行多样性预测的模型来说，它们必须对城市进行不同的分类或给出不同的预测，或两者兼而有之。这两个标准虽然并不冲突，但却很难同时满足。如果一个分类依赖于平均教育水平，而另一个分类依赖于平均收入，那么它们分类的结果可能是类似的。如果确实是这样，这两个模型可能都将是准确的，但却不是多样性的。根据每个城市名称的第一个字母创建26个类别，可以构造多样性的分类，但却很可能无法成为一个准确的模型。最重要的是，在实践中，“许多”实际上可能更接近5，而不是50。

预测的实证研究结果与这种推论一致。虽然增加模型可以提高准确性（根据多样性预测定理，必定会是这样），但是在已经拥有了一定数量的模型之后再继续增加模型，每个模型的边际贡献就会下降。例如，谷歌公司在实践中发现，仅用一位面试官评估求职者（而不是随机挑选），会使录用一名高于平均水平雇员的概率从50%提高到74%，加入第二位面试官可以把这个概率提高到81%，再加入第三位面试官则只能把这个概率进一步提高到84%，加入第四位面试官也只能提高到86%……使用20位面试官也只能将这个概率提高到90%多一点。这些证据表明，增加面试官人数的作用是有限的。

类似的结果也出现在经济学家对失业率、经济增长率和通货膨胀率进行的成千上万次的预测中。在这种情况下，我们应该把每位经济学家视为一个模型。加入一位经济学家会使预测的准确性提高大约8%，加入两位可以提高12%，加入3位可以提高15%，加入10位经济学家则能够将准确率提高大约19%。顺便说一句，假设你知道谁是最好的经济学家，那么最好的经济学家的预测只比平均水平高出大约9%。因此，3位随机选择出来的经济学家的表现就已经优于那位最好的经济学家了。[5]相信多位经济学家的平均预测、而不依赖历史上表现最好的经济学家的另一个原因是世界一直在变化。在今天的预测中表现优异的经济学家，明天就可能会泯然众人。同样的逻辑也可以解释为什么美国联邦储备系统要依赖一系列经济模型，而从来不会只依赖某一个经济模型。

这里的教益非常明确：如果能构建出多个多样性的、准确的模型，我们就可以做出准确的预测和估值，并选择正确的行动。这些定理验证了多模型思维逻辑的可靠性。但是，构建出满足这些假设的许多模型，

却不是这些定理所能做到的，也不是它们所应该做到的。在实践中，我们可能会发现我们可以构建出3个或5个很不错的模型。如果是这样，那就太好了。我们刚刚讲过，加入1个模型后可以改进8%，加入3个模型后改进幅度可以达到15%。请不要忘记，第二个和第三个模型不一定比第一个模型更好，它们也许会更糟。但是，即使它们的准确性稍差，但只要分类（字面意义）有所不同，就应该把它们加入进来。

适当的模型粒度

许多模型都能在理论上和实践中起到作用，但这并不意味着它们就一定代表正确的方法。有时，我们最好构建一个单一的大型模型。现在，我们就来分析什么情况下应该使用什么策略，同时考虑粒度问题，也就是我们应该在怎样的精细程度上划分数据。

关于应该只用一个大型模型，还是使用多个小型模型的问题，我们先回顾一下模型的7大用途：推理、解释、设计、沟通、行动、预测和探索。其中有4种用途——推理、解释、沟通和探索都要求我们进行简化。通过简化，我们可以应用逻辑来解释现象、交流思想，并探索各种各样的可能性。

回想一下孔多塞陪审团定理。在这个定理中，我们可以分析内在逻辑，解释为什么使用多模型方法更有可能产生正确的结果，也更有利于传播我们的发现。如果我们构建了一个以人格类型分类的陪审员模型，并将证据描述为语词的载体，我们就会迷失在细节的丛林中。阿根廷著名作家豪尔赫·路易斯·博尔赫斯（Jorges Luis Borges）在一篇科学论文中阐明了这一点。他描述了一批总想制作更精细地图的制图师：“制图师协会决定制作一幅国家地图，它的大小与国家大小相同，而且一对一地将土地上的每一点都标记在地图上。但是，他们的后代不像他们的祖先这样喜欢研究制图，并认为这种巨大的地图毫无用处。”

模型的另外3种用途——预测、设计和行动，却可以因高保真模型而受益。因此，如果有大数据，那么就应该利用它。根据经验，我们拥有的数据越多，模型就越精细。这一点可以通过用来梳理思维的分类模型来说明。假设我们想构建一个模型来解释数据集中的变化。为了给问题提供一个背景，不妨再假设我们从很多杂货店获取了大量数据，详细列出了数百万家庭每个月的食品支出。这些家庭的消费金额不同，我们用变差（variation）来衡量这种变化，也就是每个家庭的支出与所有家

庭的平均支出之间的差的平方和。如果每个月的平均支出是500美元，而某个特定家庭每个月的支出为520美元，那么这个家庭对总变差（total variation）的“贡献”就是400（20^2）。统计学家把一个模型中能够解释的变差比例称为该模型的R^2。

如果数据的总变差为10亿，而模型解释了其中的8亿，那么这个模型的R^2是0.8。解释的变差比例对应于模型在平均估计上的改进程度。如果某个模型估计某家庭每个月的支出为600美元，而且这个家庭的实际支出确实为每个月600美元，那么这个模型就解释了该家庭对总变差的全部贡献。如果家庭支出为800美元，但是模型的预测是700美元，那么对总变差的贡献就从原来的9万［（800-500）^{2}］，变成了1万［（800-700）^{2}］。从而模型解释了8/9的变差。

***R*²：解释变差的百分比**

$$R^2 = \frac{\sum_{x \in X}\left(V(x) - \overline{V}\right)^2 - \sum_{x \in X}\left(M(x) - V(x)\right)^2}{\sum_{x \in X}\left(V(x) - \overline{V}\right)^2}$$

其中，$V(x)$等于X中的x的值，$\overline{V}$等于平均值，$M(x)$等于模型的估值。

在这种情况下，分类模型将家庭划分为不同类别，并估计了每个类别的值。更精细的模型会创建更多的类别，而要创建这些类别就需要更多的家庭属性。如果加入了更多的类别，可以解释的变差比例就会更大。如果我们像博尔赫斯所说的那些制图师一样思考，将每个家庭都分为一类，我们就可以解释所有的变差。但是，这种解释，就像比例为1:1的地图一样，没有多大用处。

创造过多的类别会导致对数据的过度拟合，而过度拟合会破坏对未来事件的预测。假设我们想利用上个月的食品采购数据来预测本月的数据，而家庭每月的支出是会有变化的。如果一个模型将每个家庭都分为一类，那么就可以预测家庭的支出与上个月相同。由于存在月度波动，这个模型并不是一个好的预测器。通过将某个家庭与其他类似的家庭归入同一个类别中，我们可以通过对类似家庭在食品上的平均支出来构建一个更准确的预测器。

为此，我们假设每个家庭的月支出是从某个分布中抽取出来的（我们将在第5章详细讨论各种分布），再假设分布的均值和方差已知。创

建分类模型的目的是根据属性构建类别，使同一类别中的家庭具有类似的均值。如果能做到这一点，那么某个家庭在第一个月内的消费就能够告诉我们其他家庭在第二个月的支出大概是多少。当然，没有任何一种分类方法是完美的。在每个类别中，家庭的均值可能会略有不同，我们称这种情况称为分类误差（categorization error）。

构建的类别越大，分类误差就越大，因为类别越大，我们就越可能将具有不同均值的家庭集中到同一个类别中。但是，更大的类别依赖更多的数据，又可以使我们对每个类别均值的估计更加准确（参见第5章中讨论的平方根规则）。因估计均值错误而出现的误差称为估值误差（valuation error）。估值误差随类别数量的增加而减少。如果不同家庭的月支出不同，那么包含一个家庭的类别（甚至包含10个家庭的类别也一样）将无法准确估计均值，但包含1 000个家庭的类别则能够准确地估计均值。

现在，我们已经得到了关键的直觉：增加类别的数量能够通过将具有不同均值的家庭归入同一个类别减少分类误差。统计学家将这种情况称为模型偏差（model bias）。但是同时，构建更多类别则会增加对每个类别均值估计的误差，统计学家将这种情况称为均值方差的增加。因此，我们在决定要构建许多个类别时就面临着一个权衡。对于这种权衡，我们将它总结为模型误差分解定理（model error decomposition theorem），统计学家则将这个结果称为偏差-方差权衡（bias-variance trade-off）。

模型误差分解定理

偏差-方差权衡

模型误差＝分类误差＋估值误差

$$\sum_{x\in X}\left(M(x)-V(x)\right)^2=\sum_{i=1}^{n}\sum_{x\in S_i}\left(V(x)-V_i\right)^2+\sum_{i=1}^{n}(M_i-V_i)^2$$

其中，M（x）和M_i分别表示数据点x和类别S_i和V（x）的模型值，V_i表示它们的实际值。[6]

一对多

学习模型需要时间精力以及广泛的兴趣和知识。为了减少学习成

本，我们可以采用一对多的方法。我们提倡掌握适量的、比较灵活的模型，并学会创造性地应用它们。例如，我们可以使用流行病学模型来解释玉米良种的扩散、Facebook的风行、犯罪行为的传播和流行明星的“吸粉”。我们将信号传递模型应用于对广告、婚姻、孔雀羽毛和保险费的分析。我们利用进化适应的崎岖景观模型解释为什么人类不需要鲸鱼那样的喷气孔。当然，我们不能随便拿起一个模型就将它应用到任何情境之中。但是，大多数模型都是灵活的。而且，即使失败了我们也会有所获益，因为尝试创造性地使用模型能够暴露它们的局限，这是一件很有趣的事情。

一对多方法是一个相对较新的方法。过去，特定的模型只属于特定的学科。经济学家有供求模型、垄断竞争模型和经济增长模型；政治学家有选举竞争模型；生态学家有关于物种形成和复制的模型；物理学家有描述运动规律的模型，等等。所有这些模型都是针对特定目的而构建的。那个时候，科学家们不会将物理模型应用到经济学领域，也不会用经济学模型去研究大脑，就像普通人不会用缝纫机来修理泄漏的水管一样。

但是今天，将模型从各自所属的学科孤岛中“释放”出来，并将它们以一对多的方法应用到其他领域中去的做法已经取得了显著成功。经济学家保罗·萨缪尔森（Paul Samuelson）重新诠释了物理学中的模型，以解释市场如何实现均衡。经济学家安东尼·唐斯（Anthony Downs）利用经济学中描述海滩上冰激凌商店之间的竞争的模型，解释了相互竞争的政治候选人在意识形态空间上的定位。社会学家应用粒子相互作用的模型，分析不同国家的贫困陷阱、犯罪率的变化，甚至经济增长。经济学家则已经开始采用基于经济原理的自我控制模型来理解大脑的功能。[7]

一对多：更高的幂

要想创造性地应用模型，需要不断实践。为了说明“一对多”这种方法的巨大潜力，在这里以一个大家熟悉的数学公式X^N，也就是求一个变量的N次方为例，并将它作为模型应用。当幂等于2时，这个公式给出的是正方形的面积；当幂等于3时，它给出的是立方体的体积。当幂变为更高的值时，这个公式则刻画了几何膨胀或几何衰减。

超级油轮：第一个应用是考虑一艘长方体状的超级油轮，其长度是深度和宽度的8倍，表示为S。如图3-1所示，超级油轮的表面积为$34S^2$，体积则为$8S^3$。建造一艘超级油轮的成本主要取决于它的表面积，

因为这决定了所需钢材的数量。而超级油轮能够产生的收入数量则取决于它的体积。先计算一下体积与表面积之比，为$8S^3/34S^2 \approx S/4$，这表明，随着尺寸的增加，盈利能力呈线性增长。



图3-1　长方体状的超级油轮：表面积＝**34S²**，体积＝**8S³**

航运业巨头斯塔夫罗斯·尼阿科斯（Stavros　Niarchos）掌握了这个比例关系，他建造了第一艘现代超级油轮，并在第二次世界大战后的重建期间赚了数十亿美元。第二次世界大战期间使用的T2油轮长152米、深7米多、宽15米多。而现代超级油轮，例如诺克·耐维斯号（Knock Nevis），则长450多米、深20多米、宽50多米。要想象诺克·耐维斯号这样超级油轮的大小，不妨想象一下，将芝加哥的西尔斯大厦放倒，并让它漂浮在密歇根湖的水面上会是什么样子。诺克·耐维斯号大体上相当于将T2油轮放大了3倍多。然而，与T2油轮相比，诺克·耐维斯号的表面积是T2邮轮的10倍，体积则是T2邮轮的30倍。有人也许会问，那么为什么超级油轮不造得更大一些呢。答案很简单：超级邮轮必须通过苏伊士运河。事实上，诺克·耐维斯号每一次通过苏伊士运河时，都是“挤”过去的，它的两侧都只能剩下一点儿缝隙。[8]

身体质量指数：医学界通常用身体质量指数（BMI）来定义身体质量的不同类别。身体质量指数最早出现在英国，计算方法是一个人的体重与身高的平方比。[9]因此，保持身高不变，身体质量指数会随体重呈线性增长。如果一个人比身高相同的另一个人重20%，那第一个人的身体质量指数就会高出20%。

为了应用模型，我们先将人假设为近似一个完美的立方体，由脂肪、肌肉和骨骼的某种混合物构成。M表示1立方米立方体的重量。那么“人体立方体”的重量就等于它的体积乘以每立方米的重量，即$H^3 \times M$，立方体的“身体质量指数”就等于$H \times M$。到这里，这个模型还有两个缺陷：身体质量指数随身高呈线性增长；而且考虑到肌肉比脂肪更重，更健美的人会有更高的M，因此会有更高的身体质量指数。身高本应与肥胖无关，而肌肉发达本应是肥胖的对立面。即便我们使这个模型变得更加“真实”，这些缺陷仍然存在。

如果使用参数d和w来表示一个人的“深度”（前胸到后背的厚度）和“宽度”，并与高度成比例，那么身体质量指数可以写成：

$$BMI=\frac{H\times(dH)\times(wH)\times M}{H^2}=dwHM$$

。这样一来，许多NBA以及其他球类运动明星的身体质量指数将会把他们归入超重类别（BMI＞25），甚至许多世界顶尖男子十项全能运动员也不能幸免于难。[10]由于即便是身材适中、身体健康的人也可能有很高的身体质量指数，我们不应该对如下结果感到惊讶：对涉及样本总数高达数百万人的近百项研究进行的一个荟萃分析表明，体重稍稍超标的人寿命更长。[11]

代谢率：现在，应用模型来预测动物大小与代谢率之间的反比关系。每个生物体都要进行新陈代谢，也就是重复进行的一系列化学反应，分解有机物质并将之转化为能量。以卡路里计量的生物体代谢率等于维持生命所需的能量。如果我们构建小鼠和大象的立方米模型，那么从图3-2可知，小立方体的表面积与体积的比值要大得多。

我们可以把小鼠和大象建模为：身体由1立方英寸体积大小的细胞组成，每个细胞都进行新陈代谢，同时这些代谢反应产生的热量必须通过动物的体表皮肤发散掉。小鼠的表面积为14平方英寸，体积为3立方英寸，表面与体积之比约为5:1。[12]因而，对小鼠来说，每立方英寸的细胞，就有5平方英寸的体表皮肤来散热。相比之下，大象的每个发热细胞则仅有1/15平方英寸的体表皮肤来散热。这就是说，小鼠散热的速度是大象的75倍。

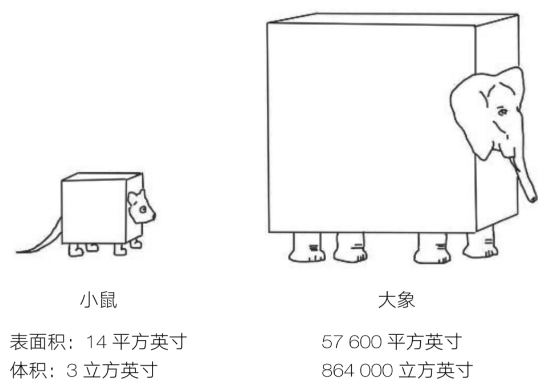


图3-2　膨胀的大象

因此，对于这两种动物来说，要想保持相同的体内温度，大象的新陈代谢就必须更慢。事实也确实如此。如果一头大象的新陈代谢速度与小鼠一样，那么这头大象每天将会需要吃下6　800千克的食物。那样的

话，大象的细胞所产生的大量热量无法完全通过它的体表皮肤发散出去。最后，大象将会热到冒烟，然后爆炸。在现实世界中，大象之所以没有爆炸，原因就在于它们的代谢率为小鼠的1/20。这个模型不能预测新陈代谢随体形大小而变化的速度，但是准确地预测了方向。更精细的模型还可以解释比例定律。[13]

女性**CEO**：最后，我们进一步增大公式中的指数，并以此来解释为什么只有较少的女性能够成为CEO。根据统计，2016年，只有不到5%的财富500强企业是由女性CEO掌管的。一个人要成为一名CEO，必须经历多次升职。我们可以将这些升职机会建模为概率事件，即一个人有一定概率可以升职。然后进一步假设，要成为CEO，必须做到每一个升职机会都不会错过。

我们假设，要成为一名CEO，至少要升职15次，这大体上相当于每两年升职一次、在30年内成为CEO。大量证据表明，升职时会出现有利于男性的“温和”的偏差。我们可以将这种偏差建模为男性升职的概率更高一些。[14]具体地说就是将这种偏差描述为男性的升职概率略高于女性的升职概率。如果将这两个概率分别设定为50%和40%，那么男性最终成为CEO的可能性几乎是女性的30倍！[15]这个模型揭示了“温和”的偏差会累积成为非常巨大的差异。10%的升职概率差异，最终变成了成为CEO可能性的30倍的差距。

这个模型也可以为如下现象提供一个新的解释：为什么女性大学校长的比例（大约25%）要比女性CEO的比例高得多？与财富500强企业相比，学院和大学的管理层级较少。一名教授只需升职3次，就可以成为大学校长：系主任、院长，然后就是校长。既然只有3个层级，那么偏差累积的程度就不会太过严重。因此，女性大学校长的比例更高，并不意味着教育机构比企业更加平等。

多模型思维

在本章的一开始，我们通过孔多塞陪审团定理和多样性预测定理为多对一的方法奠定了逻辑基础。然后，我们使用分类模型说明了模型多样性的局限性，也阐述了多个模型是怎样改进我们在预测、行动和设计等方面的能力的，同时也指出，要想构建多个不同的模型并不容易。如果可以的话，也就能达到接近完美的预测准确度了，但是我们很清楚这是不可能的。无论如何，我们的目标是尽可能多地构建有用的、多样性

的模型。

在接下来的各章中，我们将会描述一系列核心模型。这些模型突出了世界的不同部分，它们对因果关系做出了不同的假设。通过它们的多样性，这些模型创造了多模型思维的可能性。通过强调更复杂整体的不同部分，每个模型都可以发挥自己的作用，还可以成为更强大的模型集合的一部分。

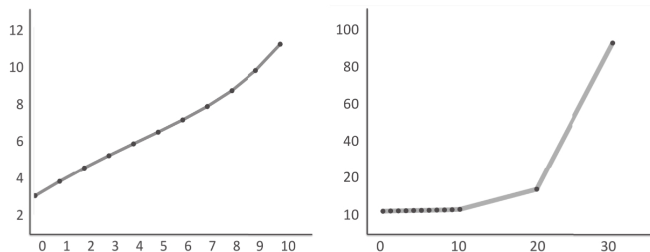
如前所述，多模型思维确实要求我们掌握多个模型，但是我们并不需要懂得非常大量的模型，只需要知道每个模型都可以应用到多个领域，但这并不容易。成功的一对多思维取决于创造性地调整假设和构建新的类比，以便将为某个特定目的而开发的模型应用到新的领域。因此，要成为一个多模型思考者，需要的不仅仅是数学能力，更需要的是创造力。这一点我们已经看得很清楚了。

装袋法与多模型

通常，我们会用模型与现有数据集中的样本拟合，然后用其余数据来检验这个模型。而在其他一些时候，我们会用模型去拟合现有数据集，然后用该模型去预测未来的数据。然而，这种构建模型的过程会产生一种张力：模型中包含的参数越多，就越能够很好地拟合数据，同时也越有可能过度拟合。好的拟合不一定意味着好的模型。

物理学家弗里曼·戴森（Freeman Dyson）曾经谈到物理学家恩利克·费米（Enrico Fermi）对他的一项研究的评论。那项研究的模型拟合度极高。“无奈之下，我问费米是不是对我们计算出来的数值与他测量出来的数值之间的高度一致性没有什么印象。他反过来问我：‘你是用多少个任意参数进行计算的？’我回忆了一下我们的截止程序，然后告诉他‘4个’。他说：‘我记得我的朋友约翰·冯·诺伊曼曾经说过，有4个参数，就可以拟合一头大象；有5个参数，就可以让大象摆动它的大鼻子了。’然后，对话就结束了。”[16]

用于“摆动大象鼻子”的估计量通常包括了更高阶的项：平方、立方，甚至四次方。高阶项的存在会带来大误差的风险，因为高阶项有很强的放大效应。10只是5的两倍，但是10^4却是5^4的16倍。下图显示了过度拟合的一个例子。

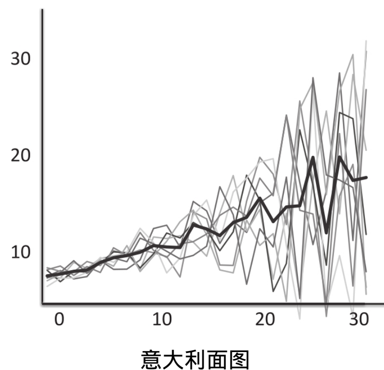


过度拟合与样本外误差

左图显示了一家生产工业用3D打印机的企业销售数据（假设），销售数据是该公司的销售团队每月（平均）上门推销次数的函数。左图显示的是一个非线性最优拟合，包括非线性项的5次方。右图则表明，如果销售团队上门推销的次数达到了30，那么该模型预测3D打印机的销售量将达到100台。如果一个客户最多只购买一台3D打印机，那这个预测就不可能是正确的。因此，由于存在过度拟合，这个模型出现了巨大的样本外误差。

为了避免过度拟合，可以避免使用高阶项。不过，一种更巧妙的解决方法是，可以采取自举聚合法（bootstrap aggregating）或装袋法（bagging）来构建模型。为了引导数据集，我们从原始数据中随机抽取若干数据点，创建多个规模相同的数据集。抽取这些数据点时，采取的是抽出后放回的方法，也就是说，在抽取了一个数据点之后，我们又将它放回到“袋子”中，下一次仍然可能会抽到它。这种技术产生了一组规模相同的数据集，每个数据集都包含某些数据点的多个副本而不包含其他数据点的副本。

然后，我们将（非线性）模型拟合到每个数据集上，以便生成多个模型。[17]这样一来，就可以把所有数据集都绘制在同一组数轴上，从而得到一幅如下所示的“意大利面图”（spaghetti graph），图中颜色最深的那条线表示不同模型的平均值。



意大利面图

装袋法能够刻画鲁棒的非线性效应，因为它们在数据的多个随机样本中都清晰可见，同时又能避免在任何单个数据集中去拟合特殊模式。通过随机样本构建多样性，然后对多个模型求平均值，装袋法很好地应用了多样性预测定理背后的逻辑。它构建了多样性的模型，如前所述，这些模型的平均值将比模型本身更加准确。

04　对人类行为者建模

时至今日，我们仍然无法说出哪一种人类行为理论是设定很成功并在各种环境下都通过了检验的。

埃莉诺·奥斯特罗姆（Elinor Ostrom）

这一章讨论本书的一个核心问题：应该怎样对人进行建模？在接下来将要给出的众多模型中，人都将成为分析的基本单元。我们将构建关于人们投票、合作、参与时尚活动、投资退休账户，以及毒品上瘾的模型。在每一个模型中，我们都必须对“人”做出假设：他们的目标是什么？他们是只关心自己的利己主义者还是利他主义者？他们可能采取的行动是什么？他们如何选择自己要采取的行动？或者说他们是否拥有选择权？

我们可以为每个模型构建任意的特殊假设，但这种做法会引起混乱并错过真正的机会。如果这样做，最终将只剩下一组特殊的构造，而且每一次要构建新模型时，都需要对人们的行为方式进行新的思考。由此产生的异质性会限制我们思考和组合模型的能力，我们将不可能成为有效率的多模型思考者。

我们遵循的方法强调一致性和多样性。或者将人建模为基于规则的行为者（rule-based　actor），或者将人建模为理性行为者（rational actor）。在基于规则的行为者集合中，我们考虑那些基于简单固定规则行事的人以及基于适应性规则行事的人。基于适应性规则行事的人能够根据信息、过去的成功或者通过观察他人的行为而改变自己的行为。正如我们在下文中将会详细讨论的那样，在这些不同的情况之间并不存在明确的界线：适应性规则有时可以解释为一个固定规则，理性行为有时也会采取简单规则的形式。

我们怎样对人建模，归根到底取决于问题的背景和想要实现的目标。我们是在预测还是在解释？是在评估政策行动吗？是在尝试设计一种制度吗？或者，是在探索？在低风险的环境中，例如要构建一个预测人们会购买什么颜色的外套或者他们会不会在看完演出后起立鼓掌的模型时，我们通常会假设人们采用固定规则。而在构建关于人们决定是不是要合作创业或信任他人的模型时，我们假设人们会学习和适应。而在高风险的环境中，我们将假设知晓相关信息的、经验丰富的人会做出最

佳选择。

在更详细地描述我们的方法之前，先来澄清一些常见的误解。许多人都是在经济学入门课程中第一次接触到描述社会现象的正式模型的，而且那些经济学模型通常依赖一个基本的理性行为者模型。在这个基本模型中，每个人都是自利的，并且有能力实现优化。这个模型通常还假设每个人都有相同的偏好和收入水平，经济学家在这些模型中求解均衡，并在此基础上评估各种冲击对市场或政策变化的影响。这些模型虽然基于不准确的假设，但却很有用。这些模型方便了经济学家之间的交流，也更有利于学生理解。

基于这种经验，许多人推断，构建正式的模型需要一种狭隘的、不切实际的“人性观”，也就是说，必须假设所有人都是自私的，而且从来不会犯错。但事实并非如此。事实上，即便是经济学家也不会这样认为。在经济学的前沿领域，早就出现了包含不完全信息和异质性行为主体的模型。在这些模型中，行为者会根据他们所了解到的东西做出调整，他们有时（尽管不总是）会关心他人的收益。当然，人们会在何种程度上表现出涉他偏好（other-regarding preferences）则取决于具体情况。例如，当向慈善机构捐款或从事志愿工作时，一个人可能会显得比在购房时更加关心他人。

尽管如此，令人遗憾的是：经济学模型总是假设自私的、不切实际的理性行为者。我们必须放弃这种观点。打个比方，如果你只是在沙滩边的海水里走了几步，那么你可能会推断海水是浅的。但是，当你游到更远的地方时，你就会开始感受到海水的深度。在这里，就让我们从近岸的浅水地带开始尝试。有时，我们会冒点儿险，进一步说明模型如何能够容纳关心他人的、有限理性的行为者。

无论做出什么假设，我们都无法摆脱假设的影响。我们被绑在逻辑一致性的“桅杆”上，不能随便制造影响。如果假设消费者的选择具有强大的社会影响力，那模型就会产生若干占据很大市场份额的产品。如果假设人们通过网络获得信息，那么填补结构漏洞的那些人将会拥有权力。

在本章的其余部分，我们先概述了在对人建模时会遇到的一些挑战：人是多样性的、易受社会影响的、容易出错的、有目的的、有适应能力且拥有自己主体性的。我们不能在一个模型中包含所有这些特征而不会产生复杂的混乱，因此我们必须做出选择。如果异质性无关紧要，那我们也许可以假设完全同质的行为主体。如果问题很简单或人们很精

明，那也许可以假设人们不会犯错误。

接下来描述理性行为者模型，并讨论其理论基础以及运用这种模型的理由（尽管它在描述层面上是不准确的）。我们的结论是，理性行为者模型是起到“黄金标准”的作用，还是“稻草人”的作用，抑或是介于这两者之间，都取决于模型的目的。理性行为者模型在预测人类行为方面的作用，不如作为沟通、评估行动和设计政策的工具那么成功。

然后，我们阐明了如何在标准的理性行为者模型中加入心理偏差和利他偏好。是否让模型包含心理偏差或（和）关心他人的偏好，仍然取决于我们正在研究的内容。某些人类心理偏差，例如损失厌恶和现世主义偏差（presentist bias）。这些假设对于退休储蓄或社会骚乱的模型可能很重要，但对于驾驶行为或疾病传播的模型可能不那么重要。

然后，我们将描述基于规则的行为。这类模型的优点是既灵活（我们可以把任何行为记下来，作为一个规则），又易处理。我们所要做的，是将这种行为用计算机程序编好码，也就是一个基于主体的模型，然后观察接下来会发生什么。这种自由当然也伴随着责任。由于我们可以选择任何一种行为规则，所以必须小心不要做出特殊的假设。在某些情况下，当给定目标函数时，可以证明所用的行为规则是一种最优行为，尽管情况并非总是如此。

最后，在本章的结束部分，我们又回过头去重新讨论理性行为作为基准行为的价值。即使人们没有优化，他们也会适应不断变化的环境和新的知识。这个观察结果带来了各种各样的难题。如果我们根据人们有心理偏差的假设，或者他们会做出不符合自身利益的行为的假设设计制度或政策，我们就不得不承担人们会改变行为的风险。你也许可以愚弄人们一次，但很难愚弄他们两次、三次。尽管不一定能得出理性是唯一合理的假设这种结论，但是逻辑确实支持将理性作为相关的基准。逻辑还支持考虑作为理性下限的简单行为规则。而且，在对任何给定情况进行建模时，我们可以应用任意数量的适应性规则和心理规则，作为探索这些极端之间巨大空间的方式。

对人建模的挑战

对人建模是一个很大的挑战，虽然模型要求低维表征，但人却是天生无法简单地加以表征的。人是多样性的、易受社会影响的、容易出错的、有目的的、有适应能力且拥有自己主体性的，也就是说，我们有行

动的能力。

相比之下，诸如碳原子和台球之类的物理对象是没有上述这6个属性的。碳原子不具备多样性，尽管它们可以在化合物中占据不同的位置，例如在丙烷中。碳原子从不违反物理定律，也不会主导有目的的生命。它们不会根据过去的经验改变自己的行为，没有主体性，也不会发起行动或转行。因此，社会科学家会时不时地讽刺：如果电子可以思考，那么物理学就会面临非常大的困难。如果电子也拥有构建模型的能力，那么物理学无疑会变得更加困难。

我们可以从多样性所带来的问题开始讨论。人们的偏好、行动能力不同，形成的社交网络、利他主义倾向以及分配给不同行动的认知资源（注意力）也有所不同。如果每个人都一样，那么建模工作就会轻松得多。有时我们会根据统计原理假设行为的多样性可以相互抵消。例如，我们可以构建一个模型，预测慈善捐赠额是收入水平的函数。对于给定的收入水平和税率，有些人可能比我们所假设的（偏好）更利他，而另一些人则可能比假设的更利己。如果偏离模型的偏差达到平均值（在第5章中，我们将给出一些能够解释为什么会是这样的分布模型），那么这个模型的预测就可能是准确的。当然，除非不同人的行动是相互独立的，否则不会出现多样性可以抵消掉的结果。在行为受到社会影响的时候，极端行为会产生溢出效应（spillovers）。当政治活动家鼓动选民时，就会发生这种情况。在下文中模拟社会骚乱时，我们会讨论多样性的这种影响。

人们所犯的错误是否能够相互抵消，取决于具体情境。认知依恋（cognitive attachment）缺失导致的误差就可能是随机且独立的，认知偏差导致的误差可能是系统性的、相关的。人们对最近发生的事件往往更加重视，并且更容易回忆起故事性的情节而不是统计数字。这类共同偏差不会被消除。

还有一个挑战与人们所渴望得到的东西有关。构建与人相关的模型时，一个主要的挑战是如何准确评估他们的目标和目的。有些人渴望财富和名声，有些人则希望为自己所在的社区乃至全世界变得更好做出贡献。在理性行为者模型中，我们直接以函数的形式表示一个人的收益。在基于规则的模型中，目的可能更加隐而不露。这是一种行为规则，人们愿意生活在一个“融合”的社区中，但如果与自己同一种族的人在社区中的比例低于10%，人们就选择离开。这样的规则显然包括人们对自己渴望得到的东西的信念。

对人建模的最后一个挑战来自人的主体性：我们有采取行动的能力，改变行为的能力以及学习的能力。也就是说，在某些情况下，人类可能是一种“习惯生物”：行动可能会超出我们的控制范围。也很少有人会主动选择沉迷于阿片类药物或贫穷。但是，归根到底是人们采取的行动产生了这些结果。

通常，当人们采取的行动产生了不好的结果时，他们会修正自己的行为。我们可以通过在模型中加入学习来捕捉这一点。人们采取的学习方式因环境而异。为了搞清楚自己需要学习多少个小时才能在考试中取得好成绩，或者自己需要每个星期锻炼多少次才能保持好身材时，人们可以根据个人经历或通过内省来学习。而在了解要到哪家杂货店购买食品，或者要不要为某个慈善项目捐款时，人们可以通过观察他人来学习。在第26章中，我们证明，在非策略性行为的环境中，学习机制一般能发挥作用，人们能够学习最好的行动。我们还将证明，在策略性的博弈环境中，则“世事难料”。而且，无论是个人学习还是社会学习，都不一定会带来好的结果。

人的这6个特征的每一个都是潜在的模型特征。如果建模时决定只包含一个特征，那么我们还必须决定在多大程度上来体现它。例如，我们要如何使演员变得多样性？需要包括多少社会影响力？人们会向他人学习吗？要如何定义目标？能拥有多大的主体性？我们所拥有的主体性（或活力）可能比自己所认为的要少。美国社会心理学家乔纳森·海特（Jonathan Haidt）(3)用骑手和大象的比喻描述了我们缺乏主体性的状态。他这样写道：“当我对自己的弱点感到惊讶时，我所想到的自我形象是一个骑在大象背上的骑手。我手中紧握着缰绳，以为自己只要动一动缰绳，就可以指挥大象，告诉它是该转弯、停步或前行。我是可以指挥它的，但只有当大象没有它自己的欲望时，我才能这么做。一旦大象自己真的想做什么事情，我根本不能左右它。”[1]我们有时候能够驾驭大象，有时候却不能。没有任何一种单一对人进行建模的方法适用于所有环境，我们不得不用多种方法对人进行建模。

理性行为者模型

理性行为者模型假设人们在给定收益或效用函数的情况下做出最优选择。这里所说的行为既可以是决策，其收益只取决于行为者个人的行为；也可以发生在博弈中，其收益取决于其他人的行为。在同时进行选

择或信息不完全的博弈中，理性行为者模型还需要设定关于其他人将会做什么的信念。

理性行为者模型

行为者个体的偏好由在一组可能的行为上定义的数学形式的效用函数或收益函数（payoff function）来表示。行为个体选择函数值最大化的行动。在博弈中，这种选择可能需要相信其他博弈参与者的行为。

我们构建了一个原始的理性行为者模型，用于描述一个人如何决定将多大比例的个人收入分配给住房支出。这个模型将行为者的效用描述为住房和所有其他消费的函数，后者包括食品、服装和娱乐。这个模型假设了住房的价格和所有其他商品的价格。当然，这个模型并不完全符合现实世界，它认为所有住房都是一样的，并将所有其他商品都归为一个名为消费品的类别，并认为它们完全等价。在这里可以暂且将这些不准确的情况放在一边不予考虑，因为这个模型的目的是解释住房支出占收入的比例。

消费的理性行为者模型

假设：行为者个体的效用来自总消费C和住房支出H，其效用函数可以写成如下形式：

$$U(C,H)=C^{\frac{2}{3}}H^{\frac{1}{3}}$$

结果：效用最大化的行为者个体会将自己收入的1/3用于住房。[2]

在这个模型中，个体用于住房的收入比例不取决于住房价格与收入水平。这两个结果在数据上都是合理近似值。[3]除了位于收入分布极端位置上的那些人之外，大多数人都将自己收入的1/3用于住房。这个发现具有重要的政策意义：如果房价下跌了10%，人们将会多购买10%的住房。这个结果也为假设同质性行为主体提供了理由。既然人们将收入的某个固定比例用于住房支出，那么住房总支出将只取决于平均收入。

使用效用函数能使模型成为可分析、可检验且易于处理的。我们可以用数据估计函数，可以推导出最优行为，还可以通过更改参数值来提

出各种各样“如果……将会怎样”的问题。假设了一个效用函数，也就意味着假设了偏好一致性，尽管那在现实世界中可能并不存在。任何偏好，要想用效用函数来表示，就必须满足某些公理。证明效用函数存在的定理需要假设一个备选方案集并确定偏好排序。想象一下，我们可以列出一个人可能购买的所有可能商品的组合，偏好排序就是这些商品组合从最受欢迎到最不受欢迎的排序。一个人可能更喜欢加牛奶的咖啡，而不怎么喜欢加柠檬的茶，如果真的是这样，那么他就将咖啡、牛奶的组合排在茶、柠檬的组合前面。

当且仅当在偏好排序中，A排在了B的前面时，效用函数赋予A的值高于赋予B的值，我们就说这个效用函数表示了偏好。偏好要与效用函数一致，就必须满足完备性、传递性、独立性和连续性。完备性要求对所有备选方案定义偏好排序。传递性排除了偏好循环，也就是说，如果有人偏好A甚于B，偏好B甚于C，那么他必定偏好A甚于C。换句话说，如果一个人在苹果与香蕉之间更喜欢苹果，在香蕉与奶酪之间更喜欢香蕉，那么他在苹果与奶酪之间必定更喜欢苹果。这个条件排除了不一致的偏好。

独立性要求人们分别评估彩票的结果。彩票是备选方案的一种概率分布，而不是备选方案本身，比如，一个可能的彩票形式是A的概率为60%，是B的概率为40%。如果，某个人对A的排序高于B，而且对于任何结果中包含了B的彩票，这个人都偏好用A代替B，那么他的偏好就满足独立性。独立性排除了过于强烈的风险规避。一个厌恶风险的人可能会在“去新奥尔良玩一趟”与“到迪斯尼世界玩一下”之间更偏好“去新奥尔良玩一趟”；同时在“肯定可以去迪斯尼世界”与“一半的机会去新奥尔良、一半的机会去迪斯尼世界”之间更偏好“肯定可以去迪斯尼世界”。

连续性要求，如果一个人偏好A甚于B、偏好B甚于C，那么必定存在这样一个彩票，以概率p得到A、以概率（1-p）得到C，他对B的偏好与对这个彩票的偏好完全一样。连续性条件排除了对某些确定结果的强烈偏好。[4]

除了人们进行优化这一可疑主张外，人们违背独立性和传递性的假设导致许多人质疑理性行为者模型的广泛使用，特别是经济学家。但是，作为建模者，我们有充分的理由采用理性行为者模型。

第一，人们往往会表现得“似乎”在最优化。他们可能会应用产生近似最优行为的规则。当人们打桌球、玩飞盘或开车时，他们当然不会写下一堆数学方程式。为了计算出接住飞盘的准确时机所需要的数学方程

式的高深程度，可能会让所有人震惊。然而，人们确实能够接住飞盘。顺便说一句，狗也接得住。因此，从人和狗接飞盘的行为来看，“似乎”两者都解决了一个困难的最优化问题。

同样的逻辑还可以扩展应用到更高维的问题上。对威斯康星州麦迪逊市大都会巴士公司（Metropolitan Bus Company）运维主管哈罗德·泽克（Harold Zurcher）的工作进行分析发现，他在是否要更换公交车引擎以及什么时候更换方面做出的决策，近乎最优。[5]虽然泽克没有用过任何数学方程式，但是他依靠启发式方法取得了成功。这些启发式来自经验，通过利用它们，泽克就能表现得“似乎”一个（几乎）完美的理性行为者了。

第二，即便人们确实会犯错，但在重复的情况下，人们的学习能力也会推动人们接近最优行为。

第三，在“赌注”（利害关系）很大的情况下，人们更应该投入足够的时间和精力来做出接近最优的选择。人们可能会为了喝一杯咖啡或买一节电池而多付30%，但他们不会在购买汽车或房子时多付30%。学习与更大的“赌注”会使人们表现得更理性，这个观点有充分的经验和实验证据支持。[6]

第四，理性行为者模型简化了分析。大多数效用函数都只有一个唯一的最优行为。一个人可以有上千种次优行为，说人们没有实现最优，就打开了一个拥有巨大可能性的盒子。如果假设人们会通过选择来维护自己的身份或捍卫文化规范，那么我们就可能无法得到一个清晰的答案。理性选择也许是不现实的，但现实主义却是以混乱为代价的。即使知道某个答案是错误的，它也可能比完全没有答案更有用，因为它至少允许我们将模型转化为数据，并讨论某些变量的变化会带来什么影响。[7]

第五，理性行为者假设保证了内部一致性。如果模型假设了次优行为且模型在公共域中，它就可以用来学习。人们可以改变自己的行为，可能不会最优化，但除了最优之外的任何假设都会受到批评，也就是不一致。我们在本章的末尾还会回到这一点。

第六，有人认为这是最重要的一个原因，也就是理性可以作为基准。[8]在设计政策、做出预测或选择行动时，我们应该考虑如果人有理性偏好并且进行最优化时会发生什么。这种做法可能帮助我们找出思维中存在的缺陷。还应该接受这样一种可能性，也就是这种做法会使我们得出这样的结论：理性行为者模型不适用，我们应该选择其他模型。有鉴

于此，我们可能会再增加第七个原因：多模型思维。如果人们应用多模型方法，犯错的可能性就会大大减少。

选择理性行为者模型的理由

“似乎”：基于智能规则做出的行为可能与最优或近似最优行为无法区分。

学习：在重复的情况下，人们应该能够接近最优行为。

大的“赌注”：在重大决策中，人们会收集信息并认真思考。

唯一性：最优行为通常是唯一的，从而使模型成为可检验的。

一致性：最优行为创建一致的模型。如果人们学会了利用这样的模型，就不会改变自己的行为。

基准：最优行为提供了一个基准，作为人们认知能力的上限。

损失厌恶和双曲贴现

理性行为者模型受到心理学家、经济学家和神经科学家的挑战。他们指出，这种模型与人类的行为方式不符。来自实验室和自然实验的经验证据表明，人在决策时会受到各种各样的偏差（包括现状偏差）的影响。我们在进行概率计算时会忽略基本比率，对确定的事情赋予的权重过高，也会表现出损失厌恶。

随着越来越多的研究者开始将行为、信念与大脑内的神经过程联系起来，硬连线偏差的证据变得非常引人注目。例如，神经经济学使用脑成像技术来研究与经济相关的行为，如对风险的态度、信心水平和对信息的反应等。[9]著名心理学家丹尼尔·卡尼曼（Daniel Kahneman）指出，到目前为止，我们已经掌握了大量支持区分两种思维方式的证据：快速、直观的基于规则的思考（快思考）和深思熟虑（慢思考）。快思考更容易受到上述各种偏差的影响。[10]从长远来看，我们可以从大脑的结构中推断出一些行为模型，但是一定要记住，大脑具有巨大的可塑性。能够通过慢思考来克服各种偏差。

此外，对于仅在少数研究中可见的任何发现，我们都应该保持谨慎。许多心理学研究的结果尚未完全得到证实。2015年一项研究表明，在主要心理学期刊上发表的100个研究结果中，有一半都无法复制。[11]更何况可复制性本身也并不意味着普遍性。而且，许多研究的被试池也没

有足够的经济和文化上的多样性。[12]如果利用更加多样性的被试池，我们应该会看到更少的行为规律，从而提供更大的理由避免对行为进行概括。

在尝试构建更符合现实的模型时，我们必须牢记易处理性这个原则。更符合现实的模型可能需要更复杂的数学。[13]这些困难或担忧当然并不意味着我们必须放弃那些心理现实行为模型，但它们确实意味着我们应该谨慎行事，并将更多的注意力放到那些已经得到很好证明的行为规律上。

下面就来讨论两种已经多次复制成功的偏差：损失厌恶和双曲贴现（hyperbolic discounting）。损失厌恶是指面对收益时，人们表现为风险厌恶，面对损失时，人们却表现为风险偏好。卡尼曼和行为科学家阿莫斯·特沃斯基（Amos Tversky）提出了一个关于这种行为的一般理论，也就是前景理论（prospect theory）。[14]损失厌恶初看上去似乎并不是非理性的，但是它意味着对于一个相同的情景，在呈现为潜在损失与潜在收益时，人们会选择不同的行为。[15]

例如，人们更偏好肯定能赢得400美元，而不怎么喜欢有机会赢得1 000美元的彩票。然而，他们却更愿意选择有可能损失1 000美元的彩票，而不愿意选择肯定会损失600美元。同样的不一致性也延伸到非货币领域。医生在收益情境时的选择是风险厌恶的，而当备选方案以损失的形式呈现给他们时，他们则愿意冒更大的风险。

前景理论：示例

收益框架：有两个备选方案。

备选方案A：肯定可以赢得400美元。

备选方案B：如果硬币正面朝上，可以赢得1 000美元；如果背面朝上，什么也得不到。

损失框架：先给你1 000美元，然后给你两个备选方案。

备选方案Â：肯定会损失600美元。

备选方案$\hat{B}$：如果硬币正面朝上，不会损失什么；如果背面朝上，你将损失1 000美元。

在这里，A和Â是等价的，B和$\hat{B}$也是等价的。根据前景理论，会有更多的人选择A和$\hat{B}$。

双曲贴现意味着，人们对近期的贴现更强。标准经济模型假设的是指数贴现（exponential discounting），也就是说，人们对未来会以恒定的贴现率贴现。对于一个年贴现率为10%的人来说，明年的1 000美元，相当于今天的900美元；而且他在未来的每一年，都会以10%的贴现率对下一年贴现。但是，大量证据表明，大多数人都不会以固定的贴现率去贴现未来。相反，他们会受即时性偏差的影响：他们对近期的贴现率远远高于更远的未来。[16]例如，如果你问人们，在从今天起20年后得到9 500美元与从今天起20年多一天后得到10 000美元之间，更愿意选择哪一个，几乎每个人都会再等一天以便多得到500美元。但是，如果你问同样一批人，在今天就可以得到9 500美元与明天才能得到10 000美元之间，更愿意选择哪一个，那么多数人都会选择现在就得到9 500美元。这就是即时性偏差的一个例子。[17]

这种偏差会导致时间不一致的行为。20年后，大多数人更愿意再等一天，以得到10 000美元。这种偏差在逻辑上并不一致。双曲贴现可以解释人们为什么会欠下巨额信用卡债务、吃不健康的食品、做出无保护措施的性行为，也可以解释许多人不能为退休进行储蓄的原因。

总之，根据对模型用途的设想，我们可以选择假设损失厌恶和双曲贴现，只要这些假设似乎更能匹配大多数人的行为。但我们也可以不这样做，主要原因是，它们可能使模型更加复杂，而不能改变我们所发现的东西的性质；或者，如果假设双曲贴现，模型可能产生不符合实际的行为。

基于规则的模型

现在讨论基于规则的模型。[18]基于最优化的模型假设人们最大化的效用函数或收益函数，而基于规则的模型则假设特定的行为。基于规则的模型可能会假设，在拍卖中，一个人的出价总是比拍卖物品的真实价值低10%；或者，如果一个人的朋友一直可以获得更高回报的话，那么这个人会“复制”这位朋友的行为。

许多人将基于最优化的模型等同于数学（模型），而将基于规则的模型等同于计算（模型）。但是基于最优化的模型和基于规则的模型之间的区别并不像人们想象的那样清晰。不妨回想一下前面给出的住房支出模型。最优行为是以一个简单规则的形式呈现的：将1/3的收入用于住房。这两种方法的关键区别在于它们的基本假设。在基于最优化的模型中，对偏好或收益的假设是最基本的；而在基于规则的模型中，对行

为的假设才是最基本的。

行为规则既可以是固定的，也可以是适应性的。固定规则意味着始终适用相同的算法。正如理性选择模型可以作为人类认知能力的上限，固定规则模型则可以作为人类认知能力的下限。在市场中，一个常见的固定规则是零智能规则（zero intelligence rule），也就是接受任何能够带来更高收益的报价。这个规则意味着永远不会采取愚蠢的（即减少效用的）行动。假设我们想要衡量单边市场机制的效率，在这种市场中，卖方对某种商品发布报价，买方要么接受、要么放弃。遵循零智能规则的卖方会随机选择一个高于该商品价值的价格，买方则会以低于该商品价值的价格购买。当我们在计算机模型中对这些行为进行编码时发现，在该市场中，零智能交易者可以得到接近完全有效的结果。因此，交易市场即便在买卖双方不理性的情况下，也可以良好地运行。[19]

而适应性规则可以在一系列行为之间切换，演变出新的行为或者复制其他行为。之所以要采取这些行动，是为了提高收益。因此，与固定规则不同，适应性规则需要效用函数或收益函数。这种方法的支持者认为，在任何情况下，只要人们倾向于采取简单而有效的规则，就应该采用基于适应性规则的模型，也就是说，既然人们以这种方式行事，那么就得按这种方式来建模。[20]虽然基于规则的模型没有对理性做出明确的假设，但适应性规则确实表现出了生态理性（ecological rationality）——更好的规则会占据主导地位。[21]

为了解释基于适应性规则的模型的工作原理，我们在这里不妨以“爱尔法鲁”（El Farol）(4)自组织协调模型为例。[22]爱尔法鲁是美国新墨西哥州圣塔菲的一家夜间营业的酒吧，每个星期二的晚上都会举办很吸引人的舞会。每个星期，都有100名潜在舞者要决定是去爱尔法鲁酒吧跳舞还是留在家里。所有这100个人都喜欢跳舞，但是如果酒吧过于拥挤，他们也就不想去了。这个模型假设了一个明确的偏好结构：一个人留在家里的收益为0；如果只有小于或等于60个人参加，那么收益为1；如果有超过60个人参加，收益为-1。

如果我们构建一个固定规则模型，那么任何结果都可能出现。假设为每个人分配这样一个规则：第一个星期，去酒吧，如果发现到场的人超过了60人，那么下一个星期就不去；再下一星期，去。那么，在爱尔法鲁酒吧，第一个星期将会涌进100个人，第二个星期却一个人都不会来，然后第三个星期又会有100个人来……与此不同，爱尔法鲁模型通过赋予每个人一组规则来创建适应性规则。每条规则都告诉个体是不是

应该去爱尔法鲁酒吧。规则有几种形式。有些是固定的规则，例如，每隔一星期去一次。其他规则是根据最近几个星期前往爱尔法鲁酒吧的人数变化趋势来制订的。例如，其中一条规则可能预测这个星期去爱尔法鲁酒吧的人数将与上个星期相同。如果上个星期到场的人数少于60人，那么这个规则就会告诉你这个星期应该去。

基于适应性规则的模型将会给每个规则分配一个分数，这个分数等于该规则给出正确建议的星期所占的百分比。然后每个人都可以采取规则集合中分数最高的那个规则。最好（分数最高）的规则将在几个星期内发生变化。对这类模型的模拟发现，如果每个人都拥有大量的规则集合，那么每个星期二都会有大约60人到场，这就是说，在没有任何中央计划者的情况下实现了协调。或者换句话说，这个适应性规则系统通过自我组织实现了几乎完全有效的结果。

爱尔法鲁模型：适应性规则

有100个人，每个人每个星期都要独立地决定是否前往爱尔法鲁酒吧。如果决定前往，且只有60个人或更少的人到场，那么这个人的收益为1，否则收益为-1，决定不前往爱尔法鲁酒吧的人收益为0。

每个人都有一套规则来决定是否参加。这些规则可以是固定的，也可以依最近一段时间以来的参加人数而定。每个星期，每个人都要按照遵循他的规则集合中曾产生过最高收益的规则行事。

我们可以在（图4-1）微观-宏观循环的框架内解释适应性规则模型（例如，爱尔法鲁模型）中的行为。在微观层面，一组个人（用a_i表示）根据规则采取行动，这些规则创建了宏观层面的现象（用$Macro_1$和$Macro_2$表示），如图4-1中向上的箭头所示。在爱尔法鲁酒吧问题中，宏观现象是过去的博弈参与者人数的序列，向下的箭头表示这些宏观现象是如何反馈到个人的行为中的。在爱尔法鲁模型中，每个人可能在应用不同的规则。如果人们所用的规则连续4个星期都导致爱尔法鲁酒吧人满为患，那么规则就会告诉人们较少参加将会带来更高的收益。当一些人转而采用这些规则后，前往爱尔法鲁酒吧的人数就会减少。微观层面的规则产生宏观层面的现象（过高的“出勤”率），后者又反馈回微观层面的规则。

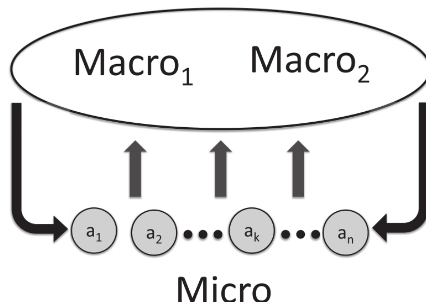


图4-1　微观-宏观循环

模型产生了什么样的结果

微观-宏观循环揭示了一个关键问题：模型中的行为主体应该有多聪明？人们有能力推断出他们行为的所有后果吗？这个循环还暗示了本书中要讨论的一个更大的问题：模型会产生什么样的结果？它是否会达成均衡？还是会产生随机性、创建一个循环，抑或导致一系列复杂的结果？

我们先从行为主体应该有多聪明这个问题开始讨论。现在不妨假设，我们认为个体都只拥有适度的认知能力，因此我们构建了一个零智能行为主体模型。个体的行为集结到一起会产生宏观层面的总体现象。如果宏观层面出现了有效或近乎有效的结果——由买方和卖方组成的单边市场就是如此，那么我们的假设就可能是合理的。既然如此易于遵循的固定规则会产生很好的结果，那么人们将几乎没有动力去开发更精致复杂的规则。

但是，当我们的模型只能产生效率低下甚至糟糕的宏观结果时，就会产生张力，爱尔法鲁模型就是这种情况。在爱尔法鲁模型中，一个共同的固定规则可能会导致完全无效率的周期性结果：这个星期爱尔法鲁酒吧是过度拥挤的，而下个星期却空无一人。面对这种效率低下的结果，我们可能更倾向认为人们会适应。他们可能会试错，可能会考虑整个形势，然后制订新的行动规则。如果将这个逻辑推向极致并假设思考成本很低，我们就会发现自己实际上已经在倡导理性行为者模型了。任何表现不佳的人都可以做得更好。虽然这是事实，但是人们也得有能力制订更好的行动规则才行。

这也就引出了一个大问题：模型产生了什么样的结果？我们有四种选择：均衡、周期、随机性或复杂性。结果的类别将决定我们对于人们

应该学会实现均衡的论点的重视程度。如果模型会在宏观层面上产生随机性，那么个人可能无法学到任何东西。好在我们的模型没有这个问题。同样的逻辑也适用于产生复杂模式的模型。在这些情况下，我们可以假设人们能够继续适应新规则，但不能假设他们可以选择最优规则。相反，宏观现象的复杂性会使最优反应显得难以置信。人们更有可能像在爱尔法鲁模型中一样，通过一系列简单的规则来应对复杂性。

产生周期或均衡的模型则可以创造一个稳定的环境，因此我们有理由期望人们可以学习，没有人会持续采取次优的行动。假设现在有这样一个交通模型：每个人都利用一个固定规则选择一条通勤路线。在这个模型中，交通系统处于某个均衡状态。假设在这个均衡中，有一个名叫莱恩的人，每天早上都要花费75分钟从卡拉巴萨斯前往洛杉矶市中心。在给定这个均衡的情况下，如果莱恩从托潘加峡谷（Topanga Canyon）抄近路走，那么她这段行程将只需要45分钟。考虑到每天能够节省30分钟的价值以及住在洛杉矶的人谈论交通的频率，莱恩确实很可能会找到这条更短的路线。事实上，她有很多种方法：可以利用地图软件的线路推荐，或者问一下自己的邻居，也可以自己多探索几次。

因此，如果模型产生了均衡（或简单的周期），并且均衡与优化行为不一致，那就意味着我们的模型存在逻辑缺陷。如果人们可以采取更好的行动，他们应该可以弄清楚，他们应该学习。需要注意的是，为了达到均衡，我们并不需要假设最优化行为。人们可以通过遵循简单的规则来产生均衡；在这种均衡中，任何人都无法通过改变自己的行为来让自己受益。在这种均衡状态下，人们看上去“似乎”正在最优化，事实上确实是。同样，这个逻辑不一定适用复杂或随机的结果。如果洛杉矶的交通模式产生一系列复杂的交通拥堵和通行减速，那么我们没有理由相信莱恩每天都能选择最优路线。事实上，她几乎肯定不能。

如果可以采取任何行动的适应性规则产生了均衡，那么这种均衡必定与致力于最优化的行为主体的行为一致。如果同样这些适应性规则产生了复杂性，那么行为主体的行为就不一定是最优的。我们还可以将这个观点表述为如下形式：最优行为可能是一种不切实际的假设，特别是在复杂情况下。另一方面，如果一个系统产生了稳定的结果，而且某个人可以采取更好的行动，那么这个人很可能会找到这种更好的选择。

同样的逻辑还延伸适用于政策干预。假设我们现在要利用数据来估计人们的行为规则，比如，一个人因为很轻微的健康问题而在午餐时间出现在医院急诊室的可能性。如果假设一个固定规则，那么我们可能会

扩建医疗设施，以保证求医的人们不必过多等待。如果人们一直持续遵循这个固定规则，就会达到新的一个均衡：中午的等待时间将会变得很短。然而，在等待时间变得更短了之后，本来不会因为扭伤了脚踝或小小的感冒就去急诊的人，也可能会决定去。这种均衡依赖于人们对次优行为的选择，例如，即使不必等待也不去急诊室。如果人们会学习，我们就不能依靠过去的数据来预测政策变化之后的结果。这个见解被称为“卢卡斯批判”（Lucas critique），是坎贝尔定律（Campbell's law）的一种变体，它指出人们对任何措施或标准的反应都会使其效率降低。[23]

卢卡斯批判

政策或环境的变化可能引起受影响者的行为反应。因此，使用过去的行为数据估计的模型将不准确。模型必须考虑到人们对政策和环境变化做出反应这一事实。

这一点我们应该已经很清楚了：对于如何对人进行建模这个问题，并不存在一个固定不变的最优答案。如何理性地制订规则及如何根据具体情况制订。我们需要的是在每种情况下尽可能做出最好的判断。考虑各种各样的不确定性，我们应该往构建更多模型的方向试错，而不是更少的模型。

即便我们倾向认为理性选择模型不切实际，也必须认识到它们的易处理性，它们所拥有的揭示激励的正确方向的能力，以及它们作为基准的价值。简单的基于规则的行为，比如零智能，也是不现实的。虽然这些假设是“错”的，但是它们仍然可以使用。它们都很容易分析，可以用来揭示给定环境下智能的重要性。

毫无疑问，人类行为发生在零智能与完全理性这两种极端情况之间，因此构建行为个体利用适应性规则的模型是有意义的。这些规则应该考虑到人们在同一个领域内的认知依恋和认知能力各不相同这一事实。因此，我们应该期待行为多样性会涌现出来，也可以期待群体内部的某种一致性。这些也都可以包含在模型当中。[24]

总而言之，考虑到对人建模所涉及的复杂性，我们有充分的理由去利用多种不同的模型。我们可能无法准确地预测人们会做什么，但是也许能够确定一系列可能性。如果可以的话，就应该多构建一些模型，我们已经从构建模型中获益，因为我们知道会发生什么。

最后，我们呼吁大家保持谦卑和同理心。在构建关于人的模型时，建模者必须非常谦卑。由于面临着多样性、社会的影响、认知错误、目的性和适应性等多种挑战，我们的模型不可避免地会出现这样那样的问题，这也正是需要采用多模型方法的原因。严谨的行为模型能够很好地拟合某些情况，并使我们能够专注于环境的其他方面。当我们拥有更多更好的数据时，更丰富的行为模型将会更合适。我们必须保持适度的期望。人是多样性的、易受社会影响的、容易出错的、有目的的、有适应能力且拥有主体性的。怎么能认为单一的人类行为模型不会出错呢？一定会。我们的目标是构建许多模型，作为一个整体，它们将是有用的。

The Model Thinker

What You Need to Know to Make Data Work for You

Part 2 模型思维

要成为一个多模型思考者，必须首先学习掌握多个模型。我们需要理解对模型的形式化描述，并知道如何应用它们。构建模型是一门艺术，只能通过不断实践才能熟练掌握。在建模中，数学和逻辑扮演着专家教练的角色，它们会纠正我们的缺漏。

05　正态分布

> 我不敢说自己比其他65个人都更聪明——但是我当然要比那65个人的平均水平更高。
>
> 理查德·费曼（Richard Feynman）

分布构成任何建模者核心知识库的一部分。从本章开始，我们将利用各种分布来构建和分析路径依赖、随机游走、马尔可夫模型，以及各种搜索模型和学习模型。如果想要度量权力、收入和财富的不平等，并进行统计检验，也需要关于分布的知识。在本书中，我们花了篇幅不大的两章专门讨论分布。本章先讨论正态分布（normal distribution），下一章讨论幂律分布（长尾分布）。我们都是从建模者而不是从统计学家的角度来讨论的。作为建模者，我们对两个主要问题感兴趣：为什么要这样看待分布？为什么分布很重要？

要解决第一个问题，就需要重新认识分布。分布以数学的方式刻画变量的变差（在某个类型内部的差异）和多样性（不同类型之间的差异），将变量表示为在数值上或类别上定义的概率分布。正态分布的形状是我们熟悉的钟形曲线形状。大多数物种的高度和重量都满足正态分布，它们围绕着均值对称分布，而且不会包含特别大或特别小的事件，例如，我们从来没有遇到过1米长的蚂蚁，也没有看到过1千克重的麋鹿。我们可以通过中心极限定理（Central Limit Theorem）来解释正态分布的普遍性。中心极限定理告诉我们，只要把随机变量加总或求其平均值，就可以期望获得正态分布。许多经验现象，特别是像销售数据或投票总数这样的总量数据，都可以写成随机事件总和的形式。

当然，并不是所有事件的规模（大小）都是正态分布的。地震、战争死亡人数和图书销量都呈长尾分布，这种分布主要由很小的事件组成，也包括极少数非常巨大的大型事件。加利福尼亚州每年都发生超过10 000次地震，但是除非你一直盯着茉莉花的花瓣看它们是否在颤动，否则你不会注意到这些地震。然而，偶然也会出现大的地震：地面裂开、高速公路塌陷，整个城市都在颤抖。

了解系统是否由于多种原因产生正态分布或长尾分布是非常重要的。例如，我们可能想了解电网是否会受到大规模停电的冲击，或者市场体系是否会产生少数亿万富翁和数十亿穷人。有了相关的分布知识，

就可以预测洪水超过堤坝的可能性、达美航空238航班准时抵达盐湖城机场的可能性，以及交通枢纽成本超过预算金额两倍的可能性。分布知识对设计也很重要。正态分布意味着不会有太大的偏差，因此飞机设计师不需要为身高5米的人预留腿部空间。对分布的理解也有利于指导行动。正如我们在下文中将会了解的那样，防止骚乱在更大程度上取决于能不能在极端情况下安抚人群，而不在于平时能不能减少不满情绪的平均水平。

在本章中，我们按结构—逻辑—功能的顺序来展开论述。我们先定义了何为正态分布，并描述它们是怎样产生的，然后回答它们为什么这么重要。我们将应用分布知识，解释为什么好的东西总是以小样本的形式出现，检验哪些效应是有显著性的，解释六西格玛（Six Sigma）过程管理为什么有效。然后回到逻辑问题，追问如果我们将随机变量相乘而不是相加会发生什么，结果是获得对数正态分布（lognormal distribution）。对数正态分布可以包括更大的事件，且均值不对称。由此，我们可以推导出，多重效应会导致更大的不平等，这个深刻的结论对提高工资的政策如何影响收入分配有重要的意义。

结构：正态分布

分布为事件或价值分配概率。每日降雨量、考试分数或身高的分布为每一个可能的结果值分配一个概率。各种统计量将分布中包含的信息压缩为单个数值，例如均值，分布的平均值。德国黑森林中树木的平均高度可能达到24米，开胸手术后的住院时间平均为5天。社会科学家经常通过均值来比较各个国家的经济和社会条件。2017年，美国的人均国内生产总值为57 000美元，远超法国的42 000美元，但是法国人的平均预期寿命则比美国人高出3年。

均值之外的第二个重要统计量是方差，可以衡量一个分布的离散程度，也就是数据与均值之间距离的平方的平均值。[1]如果分布中的每个点具有相同的值，那么方差等于零。如果一半数据的值为4，一半的值为10，那么平均来说，每个点与均值的距离为3、方差等于9。分布的标准差是另一个常用的统计量，等于方差的平方根。

可能的分布集合是无限的。我们可以在纸上任意画出一条线并将它解释为概率分布。幸运的是，我们经常遇到的分布一般都属于有限的几种类型。最常见的分布就是正态分布，也就是钟形曲线，如图5-1所示。

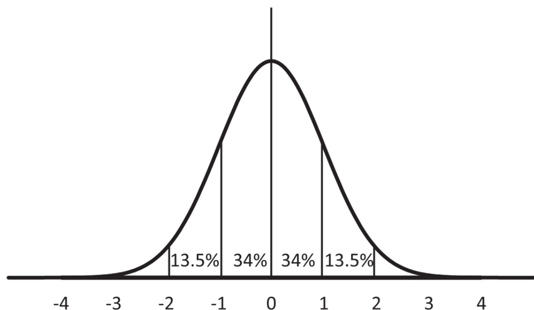


图5-1　正态分布及其标准差

正态分布的均值是对称的。如果一个正态分布的均值等于零，那么抽取到大于3的概率等于抽取到小于-3的概率。正态分布的特征在于其均值和标准差（或者等价地，其方差）。也就是说，所有正态分布的图形看上去都是相似的，大约68%的结果在均值的一个标准差内，大约95%的结果在两个标准差内，并且超过99%的结果在三个标准差内。正态分布允许任何大小的结果或事件，不过“大”事件是非常罕见的，与均值距离超过五个标准差的事件发生的概率为200万分之一。

我们可以利用正态分布的规律给各种范围的结果分配概率。如果位于美国威斯康星州密尔沃基市房子的平均面积是2 000平方英尺（1平方英尺≈0.09平方米）、标准差为500平方英尺，那么那里68%的房子面积介于1 500平方英尺到2 500平方英尺之间，95%的房子面积介于1 000平方英尺到3 000平方英尺之间。如果2019年的福特福克斯汽车平均每加仑（1加仑≈3.79升）汽油可以行驶40英里（1英里≈1.6千米），且标准差为每加仑1英里，那么超过99%的福特福克斯汽车每加仑汽油可以行驶37英里至43英里。尽管消费者希望自己的汽车越省油越好，但是一般来说不可能每加仑汽油行驶80英里。

逻辑：中心极限定理

非常多的现象都表现为正态分布：动物和植物的体型大小，学生在考试中的成绩，便利店每天的销售额，海胆的寿命，等等。中心极限定理表明为什么对随机变量求和或取均值会产生正态分布。

中心极限定理

只要各随机变量是相互独立的，每个随机变量的方差都是有限

的，且没有任何一小部分随机变量贡献了大部分变差，那$N\geq 20$个随机变量的和就近似一个正态分布。[2]

中心极限定理一个非常重要的特征是，随机变量本身不一定是正态分布的。它们可以有任何分布，只要每一个随机变量都具有有限的方差，并且它们中的任何一小部分随机变量都不贡献大部分方差。假设，在一个500人的小城镇中，人们的购买行为数据显示，每个人平均每个星期花费100美元。在这些人中，可能有些人这个星期只花50美元、下个星期则花150美元，另一部分人可能每3个星期花费300美元。而其他人则可能每个星期的花费在20至180美元之间。只要每个人的支出都只有有限的变差并且没有任何一小部分人贡献了大部分变差，那么分布的总和必定是一个正态分布，其均值为50 000美元。每个星期的总支出也将是对称的：可能高于55 000美元，也可能低于45 000美元。根据同样的逻辑，人们购买的香蕉、牛奶以及炸玉米饼的数量也都是正态分布的。

我们还可以应用中心极限定理来解释人类身高的正态分布。一个人的身高取决于基因、环境以及两者之间的相互作用。基因的贡献率可能高达80%，因此不妨假设身高只取决于基因。研究表明，至少180个基因有助于人体长高。[3]例如，一个基因可能有助于长出较长的颈部或头部，另一个基因可能有助于长出更长的胫骨。虽然基因之间存在相互作用，但我们可以假设在“长高”这件事情上，每个基因都是相互独立的。如果身高等于180个基因贡献的总和，那么身高将呈现正态分布。相同的逻辑可以证明，狼的体重和大熊猫的拇指长度也是如此。

功能：应用分布知识

我们对正态分布的第一个应用将揭示：为什么罕见结果在规模小的群体中更常见，为什么最好的学校往往规模较小，为什么癌症发病率最高的郡县人口较少。回想一下，在一个正态分布中，95%的结果位于两个标准偏差内，99%的结果位于三个标准偏差内，根据中心极限定理，一组独立随机变量的均值将是正态分布的（当然方差要满足前述要求）。由此可见，我们可以非常确信：考试分数的总体平均值也将是正态分布的。然而，随机变量平均值的标准差并不等于变量标准差的平均值，而且总和的标准差也不等于标准差的总和。相反，这些关系取决于总体大小的平方根。

平方根法则（The square root rules）

N个相互独立的随机变量，都具有标准差σ，对这些随机变量的值的标准差σ_μ和对这些随机变量总和的标准差σ_Σ，分别由以下公式给出：[4]

$$\sigma_\mu = \frac{\sigma}{\sqrt{N}} \qquad \sigma_\Sigma = \sigma\sqrt{N}$$

均值的标准差公式表明，大的总体的标准差要比小的总体的标准差低得多。由此可以推断，在小的群体中应该会观察到更多的好事和更多的坏事。事实上我们确实观察到了：最安全的居住地是小城镇，但最不安全的地方也是小城镇；肥胖率和癌症发病率最高的那些郡县的人口较少。这些事实都可以通过标准差的差异来解释。

如果不考虑样本量，直接根据离群值（异常值）推断因果关系可能会导致相当糟糕的政策行为。出自这个原因，美国统计学家霍华德·魏纳（Howard Wainer）将均值标准差公式称为“世界上最危险的方程式”。例如，在20世纪90年代，盖茨基金会和其他一些非营利机构以“最好的学校都是小学校”为依据，倡导将大学校分拆为小学校。[5]为了揭示这种推理的逻辑缺陷，试想一下，现在有两所学校，一所是只有100名学生的小学校，另一所是有1 600名学生的大学校，并假设这两所学校学生的成绩均来自相同的分布，平均分为100，标准差为80。在小学校中，平均值的标准差等于8，即学生成绩的标准差80除以学生人数的平方根10。而在大学校中，平均值的标准差则等于2。

如果以平均分为标准，把那些平均成绩在110以上的学校称为“优秀”，把平均成绩在120以上的学校称为“非常优秀”，那么将只有小学校才有可能达到这个标准。对于小学校而言，平均成绩为110时，只比总体均值高出了1.25个标准差，这类事件发生的概率大约为10%。而平均成绩为120时，则比总体均值高出了2.5个标准差，这类事件大约150所学校发生一次。对大学校进行相同的计算时，我们却会发现“优秀”阈值意味着比均值高5个标准差，而“非常优秀”阈值则比均值高10个标准差！实际上这类事件永远不会发生。因此，最好的那些学校普遍规模较小这个“事实”并不能证明小学校的表现更好。即便学校规模本身完全没有影响，“最好的学校都很小”这种事情也会发生，因为平方根法则会起作用。

检验显著性

我们还可以利用正态分布的规律来检验各种平均值的显著性差异。如果经验均值与假设均值之间的偏差了超过两个标准差，那么社会科学家就会拒绝这两种均值相同的假设。[6]现在提出这样一个假设，即巴尔的摩的通勤时间与洛杉矶的通勤时间相同。假设数据表明，巴尔的摩的通勤时间平均为33分钟，而洛杉矶为34分钟。如果这两个数据集的均值标准差都是1分钟，那么我们就不能拒绝巴尔的摩和洛杉矶两地通勤时间相同的假设。虽然二者的均值不同，但只存在1个标准差。如果洛杉矶的平均通勤时间为37分钟，那么我们就会拒绝这个假设，因为均值之间相差4个标准偏差。

但是，物理学家可能不会拒绝这样的假设，至少当数据来自物理实验时不会。物理学家采用更严格的标准，因为他们拥有更大的数据集（原子的数量远远超过了人的数量），数据也更“干净”。物理学家在2012年证明希格斯玻色子（Higgs boson）存在时所依据的证据，在700万次试验中随机出现不到一次。

美国食品药品监督管理局（FDA）所使用的药物批准程序也包含了显著性检验。如果一家制药公司声称自己研发的某种新药可以减轻湿疹的严重程度，那么这家公司就必须进行两项随机对照试验。为了构建一项随机对照试验，该公司组织了两个相同的湿疹患者群体。一组接受这种药物治疗，另一组则只使用安慰剂。试验结束后，比较平均严重程度和平均副作用发生率。然后，该公司还要进行统计检验。如果药物显著地缓解了湿疹症状（以标准差衡量）且没有显著地导致副作用，则可以批准该药物。美国食品药品监督管理局并没有使用严格的双标准差规则。治疗某种致命疾病且同时只会导致轻微副作用的药物比能够缓解真菌导致的灰指甲症状但同时却会导致骨癌发病率高于预期的药物的统计标准更低。美国食品药品监督管理局还关注统计检验的效力，也就是测试能够证明药物有效的概率。

六西格玛方法

这里要讨论的正态分布规律的最后一个应用是六西格玛方法，我们将说明正态分布是如何通过六西格玛方法为质量控制提供有效信息的。六西格玛方法是摩托罗拉公司于20世纪80年代中期提出的，目的是减少误差，该方法根据正态分布对产品属性进行建模。试想这个例子：一家企业专业生产制造门把手所用的螺栓。它生产的螺栓必须天衣无缝地与其他制造商生产的旋钮组装在一起。规格要求是螺栓直径为14毫米，但

是任何直径介于13毫米与15毫米之间的螺栓也可以接受。如果螺栓的直径呈正态分布，均值为14毫米，标准差为0.5毫米，那么任何超过两个标准差的螺栓都是不合格的。两个标准差事件发生的概率为5%，这个概率对于一家制造企业来说太高了。

六西格玛方法涉及缩减标准差的大小从而降低生产出不合格产品的可能性。各企业可以通过加强质量控制来降低误差率。2008年2月26日，星巴克超过7 000家门店停止营业3小时，目的是重新培训员工。与此类似，航空公司和医院所用的检查清单也有助于减少变差。[7]六西格玛方法降低了标准差，这样即使出现了6个标准差的误差，也可以避免出现故障。在生产螺栓这个例子中，就要求必须把螺栓直径的标准差减少至1/6毫米。而6个标准差的含义是，误差率仅为十亿分之二。实际使用的阈值假设1.5个标准差的出现是不可避免的。因此，一个六西格玛事件实际上对应于一个四个半西格玛事件，这时允许的误差率大约为三百万分之一。

在六西格玛方法中应用中心极限定理（即隐含的加性误差模型）是如此微妙，因而几乎没有什么人注意到。螺栓制造企业不可能精确地测量每个螺栓的直径，它可能会抽样几百个，并根据这样一个样本来估计均值和标准差。然后通过假设直径的变差源于多种随机效应的总和，例如机器振动、金属质量变化以及压力机温度和速度的波动，就可以利用中心极限定理推断出正态分布。这样一来，这家螺栓制造企业就可以得出一个基准标准差，然后花大力气去降低它。

对数正态分布：乘法冲击

中心极限定理要求我们对随机变量求和或求平均值，以获得正态分布。如果随机变量是不可相加而是以某种方式相互作用的，或者如果它们不是相互独立的，那么产生的分布就不一定是正态分布。事实上，一般情况下都不会是。例如，独立随机变量之间的乘积就不是正态分布，而是对数正态分布。[8]对数正态分布缺乏对称性，因为大于1的数字乘积的增长速度比它们的和的增长速度快，比如，4＋4＋4＋4＝16，但4×4×4×4＝256；而小于1的数字的乘积则比它们的和小，比如，$\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=1$，但$\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{256}$。如果将20个不均匀地分布在0到10之间的随机变量相乘，那么多次相乘后所得到的乘积将会包括一些很接近于零的结果与一些相当大的结果，从而生成如图

5-2所示的对数正态分布。

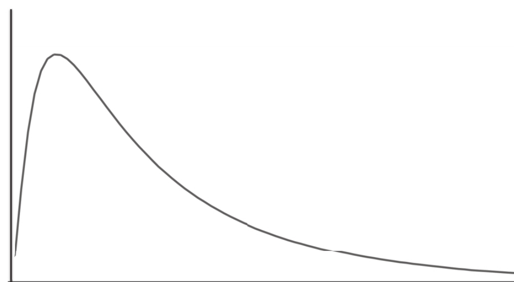


图5-2　一个对数正态分布

一个对数正态分布的尾部长度取决于随机变量相乘的方差。如果它们的方差很小，尾巴就会很短，如果方差很大，尾巴就可能会很长。如前所述，将一组很大的数相乘会产生一个非常大的数字。在各种各样的情况下都会出现对数正态分布，包括英国农场的大小，地球上的矿物质的浓度，从受到感染到症状出现的时间，等等。[9]大多数国家的收入分布也近似于对数正态分布，尽管在最顶端，许多点会偏离对数正态分布，因为高收入的人“太多”了。

一个简单的模型可以解释为什么收入分布更接近于对数正态分布而不是正态分布。这个模型将与工资增长有关的政策与这些政策所隐含的分布联系起来。大多数企业和机构都按某种百分比来分配加薪，表现高于平均水平的人能够得到更高百分比的加薪，表现低于平均水平的人则只能得到更低百分比的加薪。与这种加薪方法相反，企业和机构也可以按绝对金额来分配加薪，例如普通员工可以获得1　000美元的加薪，表现更好的人可以获得更多，而表现更差的人则只能获得更少。

百分比加薪方法与绝对金额加薪方法两者之间的区别乍一看似乎只是语义上的区别，但其实不然。[10]如果每一年的绩效都是相互独立且随机的，那么根据员工绩效按百分比加薪，就会产生一个对数正态分布。即使后来的表现相同，未来几年的收入差距也会加剧。假设一名员工因过去几年表现良好，收入水平达到了80　000美元，而另一名员工则只达到了60　000美元。在这种情况下，当这两名员工的表现同样出色并都可以获得5%的加薪时，前者能够获得4　000美元的加薪，后者却只能得到3　000美元的加薪。这就是说，尽管绩效完全相同，不平等也会导致更大的不平等。如果企业按绝对数额分配加薪，那么两名绩效相同的员工将获得相同的加薪，由此产生的收入分布将接近正态分布。

小结

在本章中，我们讨论了正态分布的结构、逻辑和功能。我们看到，正态分布可以用均值和标准差来表示。中心极限定理说明，当我们将有限方差的独立随机变量相加或求平均值时，正态分布是如何产生的。还给出了随机变量的均值与总和的标准差公式，阐述了这些性质会带来的后果。我们现在已经知道，小的群体更有可能呈现异常事件，如果缺乏对这类事件的洞察力，就会做出不正确的推断并采取不明智的行动。我们还了解到，假设随机变量服从正态分布，科学家就可以对统计检验的显著性和效力做出判断。本章还分析了在过程管理中如何利用正态假设来预测失败发生的可能性。

并不是每个量都可以写成独立随机变量的总和或平均值，因此并非所有事件都满足正态分布。有一些量是独立随机变量之间的乘积，因此它们是对数正态分布的。对数正态分布只取正值，有更长的尾巴，意味着更大的事件和更多非常小的事件。当高方差的随机变量相乘时，尾部会变得更长。长尾分布的可预测性较差，而正态分布则意味着很强的规律性。作为一个预测规则，我们当然更倾向于规律性，而不是发生很大事件的可能性。因此，如果了解了生成各种各样分布的逻辑，我们将会获益匪浅。我们可能更希望随机冲击相加，而不是相乘，以减少发生很大事件的可能性。

06　幂律分布

每个基本定律都有例外，但是你仍然需要定律，否则你所拥有的只是毫无意义的观察。那不是科学，只是做笔记。

杰弗里·韦斯特（Geoffrey West）

在本章中，我们将讨论幂律分布。幂律分布就是通常所称的长尾分布或重尾分布。在把这种分布绘制在图上时，会产生对应大事件的沿水平轴运行的长尾。例如城市人口分布、物种灭绝、万维网上的链接数量以及企业规模等，所有这些分布都有很长的尾巴，视频下载量、书籍销量、学术论文引用数量、战争中的伤亡人数、洪水和地震的分布也是如此。换句话说，在这些分布中，都包括了非常大的事件：东京有3　300万居民，J.　K.罗琳的“哈利·波特”系列畅销书卖出了5亿本，1927年密西西比河的大洪水将面积相当于西弗吉尼亚州的地区淹在了9米深的水下……[1]

要想对幂律分布与正态分布之间的巨大差异有一个直观的了解，不妨想象一下人类身高的幂律分布。如果人类身高与城市人口的幂律分布类似，而且假设所有美国人的平均身高为175厘米，那么美国人当中将会有一个人比帝国大厦还高，有超过1万人比长颈鹿还高，同时身高小于18厘米的人也将超过1.8亿人。[2]

产生幂律分布要求非独立性，通常以正反馈的形式出现。[3]图书销售、森林火灾的发生和城市人口都不同于光顾杂货店的次数，这些并不是独立的。当某个人买了一本《哈利·波特》后，其他人也可能跟着买；当一棵树着火时，火势会蔓延到邻近的树木；当一个城市的人口增加时，这个城市的基础设施会随之改善，工作机会也会随之增加，从而对其他人更具吸引力。社会学家罗伯特·默顿（Robert　Merton）把这种已经拥有更多的人未来也能够得到更多的现象称为马太效应（Matthew effect），正如《圣经》中所说：“凡有的，还要加给他，叫他有余；凡没有的，连他所有的，也要夺去。”（马太福音25:29）

既然在各种领域中都能发现发幂律分布，那么如果有某个机制可以解释所有这些幂律分布就太好了，可惜的是，这种机制并不存在。如果幂律分布的每一个实例都有一个独特的解释，那将更好，可惜的是，这也不是真的。相反，我们只拥有一系列能够生成幂律分布的不同模型，

每个模型都能解释不同的现象。

在本章中，我们将重点放在两个幂律分布模型上。第一个模型是优先连接模型（preferential attachment model），它能够解释城市规模、图书销量和网络链接等；第二个模型是自组织临界模型（self-organized criticality model），它能够解释交通拥堵、战争伤亡，以及地震、火灾和雪崩的大小等。在第12章中讨论熵时，我们还会研究第三个幂律分布模型，在那个模型中，幂律会在给定均值的条件下最大化不确定性。在第13章中，我们将证明随机游走模型中的返回次数也满足幂律分布。还有其他一些模型则表明幂律会从最优编码、随机停止规则和组合分布中产生。[4]本章还将讨论幂律分布的结构、逻辑和功能。在讨论中，我们重新评估了特别大的事件的影响，并描述在预防和规划这些事件上的能力局限。

幂律分布的结构

在幂律分布中，事件发生的概率与事件大小的某个负指数成比例。例如，我们熟悉的函数$\frac{1}{x}$就描述了一种幂律。在这个幂律分布中，一个事件的概率与其大小成反比：事件越大，发生的可能性越小。因此，在幂律分布中，小事件的数量要比大事件要多得多。

幂律分布

一个定义在区间［x_{min}，∞）上幂律分布[5]可以写成如下形式：

$$P(x)=Cx^{-a}$$

其中，指数$a>1$决定了尾部的长度，同时常数项$C=(a-1)x_{min}^{a-1}$确保总概率的分布。

幂律中指数的大小决定了大事件的可能性和大小。当指数等于2时，事件的概率与其大小的平方成比例。大小为100的事件，发生的概率与$\frac{1}{100^2}$（或一万分之一）成比例。当指数增加到3时，该事件的概率与$\frac{1}{100^3}$成比例。对于2或更小的指数，幂律分布缺乏一个可明确定义的均值。例如，从指数为1.5的幂律分布中抽取出来的数据均值永远不会收敛。换句话说，它会无限地增加。

图6-1显示了网页链接数量分布的近似图。

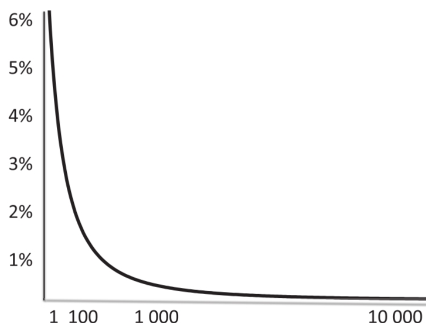


图**6-1** 网页链接的近似幂律分布

大事件的可能性将幂律分布与正态分布区分开来，因为在正态分布中，我们实际上从未见过大事件，而在幂律分布中，大事件虽然也很少见，但是它们发生的频率足以引起注意和准备。即使是百万分之一的事件也必须加以考虑。例如，地震大小的分布接近于指数大约为2的幂律。如果发生了震级大于里氏9.0级的地震，不但建筑物会被夷为平地，整个地形地貌都会变得面目全非。这是一个发生的可能性只有百万分之一的大事件，在一个世纪的时间中，这种规模的地震发生的概率为3.5%。[6]

为了更清楚地分析概率为百万分之一的大事件在正态分布与长尾分布之间的差异，现在来看一看由于恐怖袭击所造成的死亡人数的分布，它遵循幂律分布，且指数为2。[7]在长尾分布中，概率为百万分之一的恐怖袭击事件是一个差不多有800人死亡的事件。如果由于恐怖袭击造成的死亡人数满足一个均值为20、标准差为5的正态分布，那么概率为百万分之一的事件将只会导致不到50人死亡。

幂律分布有明确的定义，不是每一个长尾分布都是幂律分布。要想快速地检验某个分布是不是幂律分布，可以用双对数坐标系把该分布画出来：双对数坐标系可以将事件大小及其概率转换为相应的对数值，并将幂律分布转换为直线（图6-2）。[8]

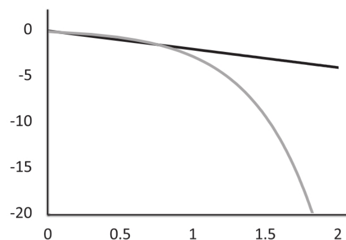


图**6-2**　双对数坐标中的幂律分布（黑色）与对数正态分布（灰色）

换句话说，在双对数坐标系中，自始至终都呈直线的图形就是幂律分布的证据，而一开始是直线然后逐渐下降的图形则与对数正态分布（或指数分布）相对应。对数正态分布图形向下弯曲的速率取决于产生分布的变量的变化。[9]当我们增大对数正态分布的方差时，对数正态分布的尾部增大，从而使在双对数坐标系中的图形更接近线性。[10]

齐普夫分布（Zipf distribution）是幂律分布的一个特例，即指数等于2的幂律分布。指数等于2的幂律分布的一个重要特征是，事件的等级排列序号乘以其概率等于常数，这个规律被称为齐普夫定律（Zipf's Law）。单词符合齐普夫定律，最常见的英语单词the出现的频率为7%，第二最常见的英语单词of出现的频率为3.5%。请注意，of的等级排列序号2乘以频率3.5%，恰恰等于7%。[11]

齐普夫定律

对于指数为2的幂律分布（$a=2$），事件的等级排列序号乘以它的大小等于常数，即：

$$事件等级\times事件大小=常数$$

包括美国在内的许多国家的城市人口分布大体上符合齐普夫定律。从美国2016年的城市人口数据可以看出，每个城市的人口排名乘以它的人口总数的值接近800万（表6-1）。

表**6-1**　城市人口分布

排名	城市	2016 年人口	排名 × 人口
1	纽约，纽约州	8 600 000	8 600 000
2	洛杉矶，加利福尼亚州	4 000 000	8 000 000
3	芝加哥，伊利诺伊州	2 700 000	8 100 000
4	休斯敦，得克萨斯州	2 300 000	9 200 000
5	菲尼克斯，亚利桑那州	1 600 000	8 000 000

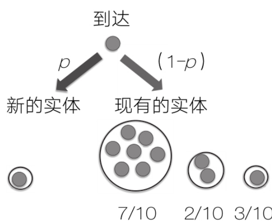
幂律分布的逻辑

现在，我们着手讨论若干产生幂律分布的模型。如果没有适当的模型，幂律分布就只是一种无法解释的模式。

我们要讨论的第一个模型是优先连接模型。模型假设实体以相对于其比例的速度增长。优先连接模型刻画了罗伯特·默顿所说的马太效应：更多导致更多。这个模型考虑了通过新移民到来而实现增长的人口。新到达的人，要么加入现有的某个实体，要么自己创建新的实体。如果是前者，那么加入现有某个实体的概率与该实体的大小成正比。

优先连接模型

一连串物体（人）一个接一个地到达。第一个到达者创建一个实体。后续每次有人到达时都应用以下规则：在概率p（较小）的情况下，新到达者创造一个新的实体；在概率（$1-p$）的情况下，新到达者加入现有的某个实体。加入某个特定实体的概率等于该实体的大小除以到目前为止所有到达者的数量。



不妨想象一下大学新生进入大学校园时的情景。第一个来到学校的学生创建了一个新的社团，第二个到达的学生以较小的概率创建了自己的社团，更有可能的是，他会加入第一个学生创建的社团。前10个到达的学生可能会创建3个社团：一个有7个成员，一个有两个成员，一个有一个成员。第11个到达的学生只会以极小的概率创建第4个社团，如果

不创建新的社团，她就加入现有的社团。如果这样做，那么她有70%的可能性加入已有7个学生的社团，有20%的可能性加入已有两个学生的社团，只有10%的可能性加入只有一个学生的社团。

优先连接模型有助于解释为什么网络链接、城市规模、企业规模、图书销量和学术引用数量的分布都是幂律分布。在这些情况下，一个行动（比如一个人购买了一本书）会增加其他人也这样做的可能性。如果从某家企业购买商品的概率与它在当前市场的份额成正比，同时如果新企业进入市场的概率较低，那么优先连接模型预测企业规模的分布将是幂律分布。同样的逻辑也适用于图书销量、音乐下载量和城市发展。

我们要讨论的第二个模型是自组织临界模型，它通过在系统中建立相互依赖关系的过程产生幂律分布，直到系统达到临界状态为止。自组织临界模型有很多种。其中，沙堆模型（sand pile model）假设有人将沙粒从距桌面几十厘米的地方洒落到桌子上。随着沙粒不断增多，一个沙堆开始形成。最终，沙子的堆积会达到临界状态，此后每加一次沙子都可能导致“沙崩”。在这种临界状态下，多加入的沙子通常要么没有影响，要么最多只会导致一些沙子下滑。这些属于幂律分布中的数量众多的小事件。但有时，只要再加入一粒沙子就会导致大规模的“沙崩”，这就是大事件。

森林火灾模型（forest fire model）也是自组织临界模型的一种。假设树木可以在一个二维网格上生长，这些树木也可能会随机地被闪电击中。当树木的密度较低时，由闪电引发的任何火灾的规模都很小，最多只会蔓延到几个格点。当树木密度变得足够高时，再被闪电击中就会导致森林大火。

森林火灾模型

“森林”最初只是一个空的$N \times N$网格。每个周期在网格上随机选择一个格点。如果该格点为空，那么就以概率g在那里种上一棵树。如果该格点上已经有树，那么闪电会以概率（$1-g$）击中该格点。如果该格点有一棵树，那么树会着火，火势会蔓延到所有连接到该格点的有树的格点。

这里需要注意的是，在森林火灾模型中，被闪电击中的概率等于1减去种树的概率。这种结构使我们能够改变种树与闪电击中树的相对速度。这是一种简化，减少了模型中参数的数量。在对各种各样的种树速

度进行试验之后，我们发现，当种树的速度接近1时，树木的密度会增加到一个临界状态：在这个相对茂密的森林中，被闪电击中有可能摧毁很大一片森林。在这种临界状态下，森林中斑块大小的分布，以及火灾大小的分布，都满足幂律分布。此外，森林还会自然而然地趋向这种密度水平。如果密度较低，密度会增加（因为火灾很小）。如果密度超过了阈值，那么任何火灾都会毁掉整个森林。因此，树木密度自组织地达到了一个临界状态。[12]

在沙堆模型和森林火灾模型中，宏观层面的变量，也就是沙堆的高度或森林的密度，都具有一个临界值。当有像沙崩或火灾这样的大事件发生时，宏观层面的变量值会减小。这两个模型的一些变体可以解释太阳耀斑、地震和交通拥堵的分布。不过，当事件发生时，不断增加的宏观层面的变量会减少，这虽然是必要的，但对于自组织临界性来说是不够的。均衡系统也具有这种特征。水通过溪流，流入和流出湖泊，但是由于水流很平稳，所以湖水的水位是逐渐变化的。通过自组织达到临界状态的关键假设是压力平稳地增加（就像水流入湖中一样），同时压力在爆发时迅速减少，这包括可能发生的大事件。

长尾分布的含义

在这里，我们讨论长尾分布的三个含义，即它们对公平、灾难和波动性的影响。根据定义，与正态分布相比，长尾分布意味着少数几个大“赢家”（大崩溃、大地震、大火灾和严重的交通拥堵）和很多的“输家”；而正态分布则是关于均值对称的。长尾分布也可能增加波动性，因为更大实体中的随机波动会产生更大的影响。

公平

如果某一个人写的书更好、创作的歌曲更有吸引力、发表的论文学术水平更高，那么他应该比其他人获得更大的名声和更多的金钱。但是，如果另一个人只是因为表现得稍微好一点，或者完全靠碰巧走运就比其他人赚到了多得多的钱、获得了大得多的名声，那就有失公平了。就像我们在优先连接模型中看到的，因为马太效应，正反馈创造了少数大赢家。在市场中，要发生正反馈，人们必须知道别人买了什么商品，而且人们必须有能力购买商品。就手机上的应用程序而言，根本不存在可能会减慢正反馈的生产限制，但是卡车就会面临这种约束。福特公司

不可能无限增加F-150卡车的产量，但是财捷集团（Intuit）却可以无限量地销售TurboTax应用程序，只要有人愿意下载。

实证研究表明，社会效应会创造更大的赢家。在音乐实验室的实验研究中，研究者让大学生挑选和下载歌曲。在第一个实验组中，被试不知道其他人下载了哪些歌曲，下载量的分布具有较短的尾部，没有出现下载量超过200次的歌曲，且下载量少于30次的歌曲也只有一首。在第二个实验组中，被试知道其他人下载了哪些歌曲，下载量的分布具有较长的尾巴，有一首歌的下载量超过300次。而且，超过一半歌曲的下载量都不到30次。

尾巴变长了，社会影响增加了不平等。如果社会影响只会导致人们下载更好的歌曲，那么这种不平等也不会造成什么问题。但事实上，这两个实验组的下载量之间的相关性并不强。我们可以将第一个实验组中每一首歌的下载次数解释为歌曲质量的一个表征，那么这项研究表明，社会影响并没有导致人们去下载更好的歌曲。大赢家的出现不是随机的，但它们其实并不一定是最好的。[13]

当然，我们必须非常小心：不能从一项研究中就得出太强的推论。然而，我们确实可以推断，卖出了5 000万册书的畅销书作家、学术论文得到了20万次引用的科学家当然是值得赞许的，但是这种极端的成功本身就表明中心极限定理是不成立的。人们不会独立地购买书籍或引用论文。惊人的成功可能意味着正反馈，也许还有一点运气。在本书的最后一章讨论收入不平等的原因时，我们还会回到这些思想上来。[14]

灾难

长尾分布还包括灾难性事件：地震、火灾、金融崩溃和交通拥堵。尽管模型无法预测地震，但确实可以深入解释为什么地震的分布会满足幂律。这些相关的知识告诉我们各种强度的地震发生的可能性。我们至少知道会发生什么，尽管不知道什么时候会发生。[15]

而且，森林火灾模型已经可以指导行动了。人们可以通过选择性地在森林中采伐一些树木来降低树木的密度，以防止大火灾的发生，也可以制造防火带。有人会说，在模型告诉我们应该这样做之前，我们早就懂得采伐树木或建造防火带了。这当然是事实，但重要的是，森林火灾模型能够让我们意识到临界密度的存在。临界密度可能因森林而异，可能取决于树木的类型、盛行风速和地形。这个模型有效地解释了为什么

森林会出现自组织临界状态。

我们还可以使用这个模型来做一个很好的类比。请回想一下，第1章中讨论了席卷整个体系的金融机构的破产，我们可以将森林火灾模型应用到那种情况下：把银行和其他金融机构想象为网格上的树，网格上的邻接则表示存在未偿还的贷款。一个银行破产相当于一棵树着火，而火势有可能会蔓延到邻近的银行。

当银行的“密度”变得越来越高的时候，这种看似浅显的森林火灾模型就预示着大规模的银行破产随时可能发生。不过，在深入探析这个类比时，我们可以发现它存在四个方面的缺点。第一，金融机构的网络并未嵌入物理空间，各家银行的连接数也不相同，有的银行可能拥有几十项金融债务，而有些银行则可能只有一两项金融债务。第二，森林中的树木不能主动采取行动来减少火势蔓延的可能性，但是银行却可以，它们可以提高自己的储备水平。

第三，一家银行拥有的连接越多，其破产会产生连锁反应的可能性就越低，因为它的损失已经分散到了更多的银行身上。例如，如果一家银行只从另一家银行借款，那么如果它在借来的1亿美元的贷款上出现了违约，第二家银行可能会破产。但是，如果第一家银行是从其他25家银行分别借款的，那么任何一家银行都不至于受到重创。在这种情况下，银行体系可以很好地消化这个违约事件而不会崩溃。[16]

第四，从一家银行的破产到另一家银行的破产，这种蔓延会不会出现还取决于银行的投资组合。如果两家“相连”的银行拥有相似的投资组合，那么当一家银行破产时，另一家银行也可能早就脆弱不堪了，这时银行破产蔓延的可能性就很大。如果整个网络中的所有银行都拥有相同的投资组合，那么最糟糕的情况就很可能会出现。在这种情况下，当一家银行破产时，就可能会出现普遍的银行破产。[17]但是，如果每家银行分别持有不同的投资组合，那么一家银行表现不佳并不意味着其他银行也表现不佳。在这种情况下，银行破产就可能不会蔓延。因此，一个模型要想真正有用，就必须考虑到各种不同的投资组合。如果没有这些信息，那么即便知道哪些银行对其他银行负有未偿还债务也不足以预测或防止银行破产，而且银行之间的高互连性的净效应也是不明确的。

波动性

最后，我们讨论最微妙的一个含义。如果组成幂律分布的实体规模

出现了波动，那么幂律的指数就可以作为衡量系统层面波动性的一个代表。由此可以推断，企业规模的分布应该会影响市场波动性。例如，我们可以将某个国家的国内生产总值视为数千家企业的总产量。如果各家企业的生产水平相互独立且变差有限，那么根据中心极限定理，这个国家的国内生产总值分布将服从正态分布。也就是说，企业生产水平的差异越大，总体波动性就越大。如果企业规模的长尾分布导致生产水平上更大的变差，那么这种长尾分布也必定与更大的总体波动性相关。

对美国波动性模式的实证研究表明，波动性在20世纪70年代和80年代有所上升，然后在接下来的20年间又下降了，有人将后面这20年称为“大稳健”（Great Moderation）。[18]但是，从2000年前后开始，波动性再次上升。研究显示，可以通过企业规模分布的变化来解释这种波动性演变的模式。[19]随着企业规模分布的尾部变得越来越长（越来越短），最大的企业对波动性的影响越来越大（越来越小）。换句话说，总体波动性会随企业规模分布的尾部变长（变短）而增加（减少）。1995年，当总体波动性较低时，沃尔玛的营业收入为900亿美元，相当于美国国内生产总值的1.2%。到了2016年，沃尔玛的营业收入增加到了4 800亿美元，占国内生产总值中的百分比提高到了2.6%，沃尔玛在美国国内生产总值中所占的份额增加了一倍多。2016年，沃尔玛收入的增加或减少可能导致的总体波动性因而也增加了一倍多。

没有人能够反驳这个观点的逻辑。因此，相关的问题是，一个经过校准的模型到底能不能生成与实际波动水平相对应的效应。校准拟合结果表明，确实非常接近。企业规模的分布很好地对应于大稳健时期的历史证据。虽然这种相关性并不能证明是企业规模分布的变化（而不是政府对经济的有效管理或更好的库存控制）导致了这种变化，但是它确实足以“阻止”我们拒绝这个模型。[20]这些证据还为我们在未来评估各种波动时将这种模型加入我们的工具箱提供了很好的理由。

设想长尾分布的世界

在长尾分布中，大事件发生的概率必须加以考虑。在本书讨论的多个模型中，长尾分布是由于反馈和相互依赖性而产生的。我们应该高度注意这个结果。随着世界中相互联系性的提高和反馈的增加，我们应该会观察到更多的长尾分布，同时现在关注的这些长尾分布的尾部也可能会进一步拉长。这就是说，不平等可能会增加，灾难可能变得更大，波

动性也会变得更加剧烈。这些都是不可取的。

到目前为止，我们都是在宏观层面上讨论这些事件的可能性的。它们也同样可能发生在更小的尺度上。波士顿的中央隧道工程（Big Dig）是一条穿过市中心的长达5 000多米的隧道，它是一个中等规模的灾难的典型例子。这个项目花费了140亿美元（相当于最初预算的3倍多），并成了美国有史以来最昂贵的公路项目。根据模型思维的方法，我们不会把这个项目简单地视为一个单独的项目，而是作为很多子项目的总和：挖掘深沟、浇筑混凝土隧道、设计排水系统、建造墙壁和“顶盖”。项目的总成本等于各个子项目成本的总和。

如果每个子项目的成本都是相加的，那么这个项目的成本分布将是正态分布。[21]然而，各个子项目的成本是相互关联的。原本计划用来将顶盖黏合到位的那种环氧树脂强度不够时，就不得不用成本更高、强度更大的另一种环氧树脂来代替，从而增加了项目的成本。而且，第一种环氧树脂的失效还产生了移除和更换折叠顶盖的额外成本。这些工作反过来又需要重做项目的其他几个部分。于是总体成本增加了一倍以上，因为每个项目必须撤销然后重做。这种相互依赖性最终导致了一个大型且昂贵的事件。

大事件发生的可能性使计划变得非常困难。像地震这样自然灾害的分布符合幂律分布。因此，大多数事件都是很小的事件，但是有些事件一旦发生就会很大。如果灾难性事件遵循指数为2的幂律分布，那么政府就必须时刻保留大量的储备金或者至少做好应对的准备，必须未雨绸缪。如果政府为了这个目的而在应急基金账户中保持了巨额盈余，那么如果没有大事件发生，政府可能会阻止自己花掉这笔钱或减税。

搜索与机会

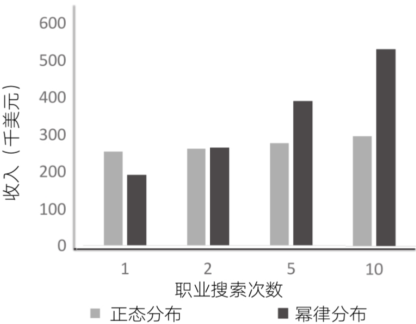
我们可以在某些搜索模型中应用关于分布的知识来解释为什么一个人获得机会的数量可能与他的成功经历密切相关。在这里，事实上是将一类模型（分布模型）嵌入了另一类模型（搜索模型）。我们在搜索的时候，无论是搜索新鞋、工作职位还是度假胜地，其实是不知道所选择的价值的，直到去真的尝试它。不过，我们可能会对所选择的价值的分布有所了解，例如它的均值、标准差，以及这种分布是正态分布还是长尾分布等。

在这里，我们将职业选择建模为一个搜索过程。给定某个行

业，一个人尝试某条职业道路。我们将这种行为建模为从一个分布中抽取某个事件。假设，这个人可以坚持这个职业选择或再试一次，再试一次对应于从分布中的另一次抽取。例如，考虑一个有才华的年轻科学家的职业选择。她可以选择去医学院深造，也可以选择去研究量子计算。医学院提供了一条更安全的道路，选择研究量子计算则可能成为一名创业企业家并承担更多风险。为了解释这些差异，我们将医生的工资分布表示为均值25万美元，标准差25 000美元的正态分布，并把量子计算企业家的工资分布表示为指数为3、期望工资为20万美元的幂律分布。[22]

再假设，在每个行业内，这位科学家也可以尝试多种职业。也就是说，她可以搜索。医生可以从肿瘤科转入放射科，一个企业家破产后也可以重整旗鼓继续尝试创业。但是每一次转换职业都要付出一定的成本：对于一个医生来说，这意味着要接受更多的培训；对于一个从事量子计算的企业家来说，这意味着需要付出更多的时间去从事没有报酬的工作。

另外再假设，这位科学家认为这两个职业同样有意思，并且会根据薪资水平来做出选择。我们的模型证明，哪种选择更好取决于有多少次尝试新职业的机会。如果她必须坚持自己的第一个职业选择，那么成为一名医生就可以获得更高的期望工资；如果她有足够的资源持续尝试，努力成为一名企业家，那么最终她将从长尾中获得高薪。假设在每个职业中分别进行1次、2次、5次和10次职业搜索，下图显示在20次测试中得到的平均最高工资。如果这位科学家有机会在量子计算初创企业中尝试10次，那么她的薪资将会是她选择进入医学院深造并尝试10个职业后收入的两倍。



如果拥有的财富和家庭的支持与一个人不得不尝试新职业的机会数量相关，那么模型的预测是，更富有的人将选择风险较高的职

业。[23]专利证据与我们的模型是一致的。一个人成功申请专利的可能性与他的数学技能相关，数学能力排名前1%的人更容易获得专利。而且在这数学能力排名前1%的人当中，收入排名前10%的家庭的人更有可能拥有专利。[24]至少有两个模型可以解释这种差异：一个模型假设更贫穷且有才华的学生没有上大学的机会，他们可能正在从事日常工作，从未有机会在进入医学院深造还是去研究量子计算之间做出选择，另一个模型则假设更贫穷的学生会选择更安全的职业。

机会的增加可以创造风险激励，这个逻辑可以应用到很多领域。风险资本家经常冒险，因为他们有机会进行多项投资。只要投中了一个独角兽（市值10亿美元以上的公司），不仅可以补偿多次失败的投资，还可以带来很大的利润。研究药物的实验室也愿意承担风险，花费数十亿美元用于药物开发。甚至在决定午餐吃什么时，我们也可以应用同样的逻辑。长途旅行并在某个不熟悉的小镇短暂停留时，我们一般更喜欢选择连锁餐厅用餐；但是，如果真的搬到那个小镇去居住，我们就会尝试多家餐厅。

07　线性模型

是的，我承认我在说谎。但为什么你非要强迫我给出一个线性解释呢！线性解释几乎总是谎言。

埃莱娜·费兰特（Elena Ferrante）

模型通常假定变量之间存在某种特定的函数关系。这种关系可以是线性的，也可以是非线性的，或者可以包括阈值效应。在这些模型中，线性模型是最简单且应用最广泛的。本章的重点就是线性模型。教育对收入的影响、因锻炼而增加的期望寿命，以及收入对选民投票率的影响，都可以用线性模型来解释。

在本章的开头部分，先回顾一下单变量线性函数。然后讲解了如何通过回归将数据与线性函数拟合，并揭示各种效应的符号、大小和显著性。我们还讨论了为什么误差、噪声和异质性意味着数据不会全部落在回归线上。接着，我们扩展了线性模型以容纳更多的变量，并讨论了如何拟合多元线性模型。为了建立多变量模型的直觉，我们将成功建模为技能和运气的线性函数。本章的结尾部分总结了如何依靠数据和回归指导行动、减少错误，但是这样做也可能会导致边际行为，进而导致保守的行为。确实，“唯大系数论”思维可能会扼杀创新。为了确定更多的创新项，我们可能需要考虑构建其他更具推测性的模型。

线性模型

在线性关系中，由于第二个变量的变化而导致的第一个变量的变化量不依赖于第二个变量的值。假设树木的高度与树木的年龄呈线性关系，那么树木每年生长的高度相同。假设房子的价值随它的面积（平方米）线性增加，那么房子面积扩大200平方米所带来的房子价值的增量，等于房子面积扩大100平方米所带来的房子价值增量的两倍，400平方米的扩大使房子的价值增加了4倍。

线性模型

在线性模型中，自变量x的变化，会导致因变量y的线性变化，用如下方程表示：

$$y = mx + b$$

其中，m等于直线的斜率，b等于截距，即当自变量等于0时的因变量值。

线性回归模型的目标是找到能够最小化到各数据点的直线。线性回归可以解释犯罪、洗衣机销量，甚至可以解释葡萄酒价格的变化。[1]假设我们找到了一组年龄介于20岁到60岁之间的成人的年龄数据以及他们每个星期走路的距离，可以发现如下回归方程：

$$\text{第}i\text{个人步行的英里数} = -0.1 \times \text{年龄}_i + 12 + \varepsilon_i$$

这个回归方程不仅告诉我们这种效应的符号（距离随年龄的增长而减少），还告诉我们这种效应影响的大小（年龄每增加1岁，距离减少1/10英里）。在这个例子中，截距并不重要，因为它位于数据范围之外，也就是说，数据原本就不包括年龄接近于0的人。根据这个方程，我们可以预测，一个40岁的人每个星期步行8英里，而50岁的人则每个星期步行7英里。但是，用于产生回归的数据不会全部都落在回归线上。

图7-1显示了用于生成回归线的假想数据。其中灰色圆圈代表的人名叫博比，他40岁了，每个星期步行11英里，比模型估计的要多走3英里。为了使数据与模型一致，我们在方程中给每个数据点增加了一个误差项，用ε表示，等于模型估计值与因变量实际值之间的差异。博比的误差项等于+3英里。

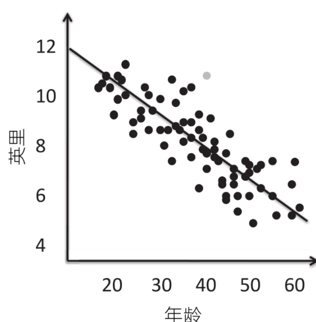


图7-1 散点图和回归线

在社会环境和生物环境中，我们不能期待会有完美的线性拟合。结果通常取决于多个变量，但是根据定义，单变量回归只能包含一个变量。由于这些被省略的变量，预测值可能会偏离实际值。博比走的路可

能要比预期更多，因为作为植物学教授，他要带他的学生到树林里采集标本。这个模型没有把职业作为一个变量，但是职业有助于解释为什么图7-1中的数据点没有落在回归线上。ε项也可能由测量误差引起。如果人们忘记随身携带智能手机，或者将智能手机借给他人，那么利用智能手机收集的健身数据就会包含误差。此外，环境噪音也可能导致误差，比如人们可能会因为坐在颠簸的汽车上而获得额外的步行里程。[2]

回归线越靠近数据，模型解释的数据越多，R^2就越大（得到解释的百分比越大）。如果数据全部都恰好位于回归线上，R^2就等于100%。

符号、显著性和大小

线性回归可以告诉我们关于自变量系数的如下内容：

符号：自变量与因变量之间的正相关或负相关。

显著性（p值）：系数上非零符号的概率。

大小：对自变量系数的最佳估计。

在单变量回归中，回归线与数据拟合得越好，我们对系数的符号和大小就越有信心。统计学家使用p值来表示系数的显著性，p值等于基于回归的系数不为零的概率。p值为5%意味着数据由一个系数等于零的过程生成的概率为1/20。显著性的标准阈值是5%（通常用*表示）和1%（通常用**表示）。但是，显著性并不是我们唯一关心的东西。一个系数可能是显著的，但是却很小。如果真的是这样，就可以对相关关系很有信心，但是变量的影响其实不大。又或者，也可能系数虽然不显著但却很大，这通常发生在有噪声数据或数据带有许多遗漏变量的情况下。

为了阐明如何利用回归来指导行动，不妨想象一下这样一家销售香料的公司。该公司供应超过100种香料。客户会购买包含6种、12种或24种香料的包装。客户下单后，员工负责包装和运输。将每8小时的班次的订单数量作为员工工作年限的函数进行回归，结果如下：

$$完成的订单数 = 200 + 20^{**} \times 工作年数$$

在上面的方程中，工作年数前面的系数20的显著性水平为1%。我们可以确信它是正的。如果这种关系是因果关系，那么这个模型就可以用来预测每个员工每个班次可以完成的订单数量（作为工作年数的函数），还可以使用这个模型来预测某个在职员工明年可以完成的订单数

量。在这里，有一个模型的实例，既可以给出预测，也可以指导行动。

相关关系vs. 因果关系

回归所揭示的是变量之间的相关关系，而不是因果关系。[3]如果先构建了某个模型，然后用回归检验模型的结果是否得到数据的支持，但那也不能证明因果关系。但是，在我们能够用回归发现显著的相关性之前，有一种方法远比回归方法好，这种方法就是通常所称的“数据挖掘”（data mining）。但是，数据挖掘存在识别与其他因果变量相关的某个变量的风险。例如，数据挖掘可能会揭示维生素D的水平与身体总体健康程度之间存在显著的正相关关系。人们多晒阳光有利于吸收维生素D，因此这种效应可以归于生活方式更积极的那些人在户外度过的时间更长，从而健康状况更好。或者回归可能会发现，某个大学的学术表现与参加马术队的学生人数存在显著相关。但是，马术队与学术水平之间可能并不存在直接的因果关系，但它们与平均家庭收入和学校资助水平相关。

数据挖掘还可能导致虚假的相关关系，即两个变量只是偶然相关。我们可能会发现，名字较长的公司可以获得更高的利润，或者居住在比萨店附近的人更容易患流感。事实上，使用5%的显著性水平阈值，每检验20个变量就会发现有一个是显著的。因此，如果尝试足够的变量，肯定会发现某些显著但虚假的相关性。

我们可以通过创建训练集（training set）和检验集（testing set）来避免报告虚假相关。在训练集上发现的相关性，如果也存在于检验集上，就更可能是真实的。但即便是这样，我们仍然无法保证那就是因果关系。为了证明因果关系，还需要进行一个实验来操纵自变量并观察因变量是否会随之发生变化，或者也可以想办法找到可以证明这类因果关系的自然实验。

多元线性模型

大多数现象都有不止一个因果变量和相关变量。一个人的幸福可以归因于身体健康、婚姻美满、子女、宗教信仰和财富等。一栋房子的价值取决于室内面积、庭院大小、浴室数量、卧室数量、建筑类型以及当地学校的质量等。在解释房子价值的时候，可以把所有这些变量都包含在回归中。但是必须记住，随着添加更多的变量，也就需要更多的数

据，不然无法得到显著的系数。

实力-运气方程

在讨论多元回归之前，先引入迈克尔·莫布森（Michael Mauboussin）的实力-运气方程，以便对多元方程有一个直观的认识。[4] 这个方程说的是，任何成功，无论是日常工作中的成功、体育运动上的成功，还是游戏时的成功，都可以视为实力-运气的一个加权线性函数。

实力-运气方程

$$成功=a\times 实力+（1-a）\times 运气$$

其中，a位于区间［0，1］上，是技能的相对权重。

如果给实力和运气分配适当的权重（也许通过利用现有数据进行回归，可以得到这样的权重），我们就能够运用这个模型来预测结果。例如，假设一家休闲汽车销售公司的经理发现，用销售数量来衡量的成功有很大的运气成分，那么他就会期待回归均值：本月取得了很好业绩的销售人员下个月可能会回到平均水平。然后，这个经理就可以利用这个模型来指导行动了。比如，他不会为了争取一个连续两个月都取得非常不错业绩的销售人员而付出比竞争对手高很多的薪资。相反，如果回归表明运气对成功几乎没有任何作用，那么连续两个月的业绩就可以作为未来业绩表现的一个很好的预测器。在这种情况下，经理就应该为这个最佳销售人员提供有竞争力的报酬。[5]

同样的逻辑也适用于CEO薪酬的决定。在那些“运气决定了成功”的行业中，董事会不应该向CEO发放高额奖金。石油公司的利润取决于原油的市场价格，那是一个公司无法控制的变量。因此，一家石油公司的董事会不应该因为某一年公司业绩不错就给CEO发放巨额奖金。相反，广告公司则不然：如果广告公司业绩表现良好，那么给CEO发放巨额奖金就是一件明智的事情。简而言之，要奖励实力，而不要为运气去买单。事实上，那些很成功的公司都不会为运气付出太多。

即便是最简单的模型，例如上面这个实力-运气方程，也能帮助我们得出深刻的见解。进一步思考这个方程可以发现，即便是在那些成功几乎完全取决于实力的环境中，例如跑步、骑自行车、游泳、下棋或网球比赛，如果不同的参赛者之间实力差异很小，那么运气就会在很大程度

上决定谁输谁赢。我们可以预期，在竞争最激烈的比赛中，比如奥运会，进入决赛的选手之间的实力差异很小，因而运气就非常重要了。莫布森把这种情况称为“实力悖论”（paradox of skill）。

历史上最伟大的运动员之一迈克尔·菲尔普斯（Micheal Phelpls）可以说同时位于这个悖论的两端。在2008年奥运会的一场决赛中，菲尔普斯在100米蝶泳快结束时仍然落后于米洛拉德·卡维奇（Milorad Cavic）。然而幸运女神眷顾了他，菲尔普斯率先触到了池壁。然而，在2012年奥运会的一场决赛中，菲尔普斯一直领先于查德·勒·克洛斯（Chad le Clos），但是幸运女神这次没有眷顾他，勒·克洛斯率先触到了池壁。菲尔普斯拥有令人难以置信的实力，但是上一次胜利和这一次失败，却都是运气的产物。

多元线性回归

多元线性回归模型拟合了具有多变量的线性方程，当然同样要最小化到数据的总距离。这些方程包括每个自变量的系数。下面的方程反映了这样一个假设的回归输出：学生在数学考试中的成绩，是学生学习的小时数（HRS）、学生家庭社会经济状况（SES）和上“快班”课程的数量（AC）的函数。

$$\text{数学成绩} = 21.1 + 9.2^{**} \times \text{HRS} + 0.8 \times \text{SES} + 6.9^{*} \times \text{AC}$$

根据回归分析的结果，学生每多学习一个小时，数学成绩会提高9.2分。这个系数有两个*号，因此它在1%的水平上显著，这意味着很强的相关性，尽管不是因果关系。这个方程也表明，每参加一个“快班”课程，数学成绩能够提高近7分，这个系数也是显著的，但仅仅在5%的水平上显著。家庭社会经济状况这个变量的取值为从1（低）到5（高），系数也为正，但是与零没有显著差异，因此我们可以认为它可能没有什么因果关系。

有了这样一个（或任何形式的）回归方程，我们就可以预测结果。这个模型预测，如果花7个小时学习，并同时参加一个“快班”课程，数学成绩就能够达到90分左右。这个模型还可以用来指导行动，但必须保持谨慎，因为我们无法推断因果关系。数据表明，花时间学习和参加“快班”课程的学生成绩更好。但是，花时间学习和参加“快班”课程这两个因素也可能没有什么用，因为也许存在选择性偏差（selection bias），那些花更多时间学习、参加“快班”课程的学生，数学成绩可能本来就更

好。

即便回归不能说明是什么原因导致数据呈现出来的特定模式，但是至少可以排除其他解释。以美国种族之间的巨大财富差距为例：2016年，白人家庭的平均财富（约11万美元）是非洲裔美国人家庭和拉美裔美国人家庭的10倍。各种各样的原因都可以用来解释这种差距，包括制度因素、收入差距、储蓄行为差异或结婚率差距等。回归可以为其中一些解释提供支持并排除其他解释。例如，回归分析表明，非洲裔美国人的婚姻状况与家庭财富之间没有显著关系，因此婚姻状况不能成为这种财富差异的原因。此外，收入差距虽然相当大，但是也不足以解释这种财富差距。[6]

大系数与新现实

如前所述，线性回归模型在科学研究、政策分析和战略决策中都发挥着重要作用，部分原因是因为线性回归模型容易估计和解释。而且，随着数据可得性的不断改善，线性回归模型得到了更广泛的应用。“要信只信上帝，要认只认数据”（In God we trust. Everyone else must bring data.）这句话在商界和政界都可以经常听到。对数据的这种依赖（通常意味着线性回归模型），可能会导致我们过于倾向边际行动（marginal action），远离重要的新思想。企业、政府或基金会，都致力于收集数据，拟合线性回归模型，试图找到有最高统计显著性系数的变量，这种努力几乎肯定会导致调整该变量并获得边际收益的行为。

在采取行动的时候，最好选择具有较大系数的变量，而不要选择具有较小系数的变量。与此同时，“大系数至上”这个思路建立在“保守主义”的基础上，它会使我们将注意力集中到较小的改进上，而无法再关注全新的政策。“大系数至上”的另一个问题是大系数的大小对应于给定现有数据的边际效应。正如我们在下一章中会阐述的，通常这种效应将会随着变量值的增大而减少。如果确实是这样，那么当我们试图利用它时，大系数就会变小。

大系数与新现实

线性回归揭示了自变量与我们感兴趣的（因）变量之间的相关程度。如果这种相关是因果关系，那么具有大系数变量的变化就会产生很大的影响。基于大系数的政策在保证能够带来改进的同时，

排除了涉及更多根本性变化的新现实。

“大系数至上”思维方式的替代者是“新现实思维”。如果说，大系数思维可以拓宽道路、建造高利用率的车道以减少交通拥堵，那么新现实思维就相当于建造了铁路和公共汽车系统。大系数思维为低收入家庭的学生购买计算机提供补贴，新现实思维则直接为每个人都提供了计算机。大系数思维改变了飞机上座位的宽度，新现实思维则创造了一个使用可互换吊舱的飞机机舱。大系数思维已经相当不错了，因为基于证据的行为是明智的，但我们也必须同时关注重要的新思想。当我们遇到重要的新思想时，可以用模型去探究它们是否可行。对青少年交通事故的回归也许会告诉我们，年龄的系数是最大的，这意味着提高驾驶年龄的政策也许能起到一定作用。还可以采取更多的政策，例如禁止夜间驾驶的宵禁、通过智能手机自动监控青少年驾驶员或限制青少年驾驶汽车的乘客数量等。这些新现实政策带来的效果可能比大系数带来的要好。

小结

总而言之，线性模型需要假定效应大小不变。线性回归为我们对数据进行第一轮加工提供了一个强大的工具，有了它，我们能够识别出变量的符号、大小和显著性。如果我们希望了解咖啡、酒精或苏打水对健康的影响，就可以进行回归分析。我们可能会发现，喝咖啡会降低心血管疾病的风险，适量饮酒也有同样的效果。这也就是说，在现有数据范围之外推断线性效应时必须非常小心。我们绝对不能推断，喝30杯咖啡、6瓶葡萄酒会是个好主意。我们不应该用线性模型对过于久远的未来进行预测。从1880年到1960年，加利福尼亚州的人口增长率为45%，如果进行线性预测，那么我们将会预测2018年加利福尼亚州的人口会达到1亿人，但这超出了实际人口水平的两倍。

请记住，线性模型只是一个开始，大多数有趣的现象都不是线性的。因此，回归模型通常会包括非线性项，例如年龄的平方、年龄的平方根，甚至包括年龄的对数。为了解释非线性，我们还可以将线性模型首尾相连，这些连接起来的线性模型可能近似于曲线，就像我们可以使用直边的砖块来砌出弯曲的路径一样。虽然线性可能是一个强大而不切实际的假设，但是它至少提供了一个很好的起点。在给定了数据的情况下，可以使用线性模型来检验我们的直觉判断。然后，我们可以构建更精细的模型，其中变量的影响会随着它的增加（收益递减）或变得更强

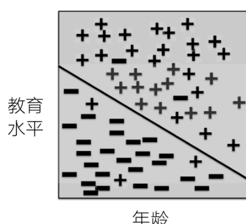
大（正回报）而减弱。这些非线性模型正是下一章要研究的重点。

对数据的二元分类

在当今这个大数据时代，组织普遍使用根据模型建立的算法对数据进行分类。政党可能想要了解哪些人投了谁的票，航空公司可能想知道常客的特点，某项活动的组织者可能想要了解哪些人会参加这项活动。在所有这些情况下，它们所使用的方法都将相关的人分成了两组：一组是“正”的（＋），也就是购买了、贡献了、注册了的人；另一组是“负”的（-）。

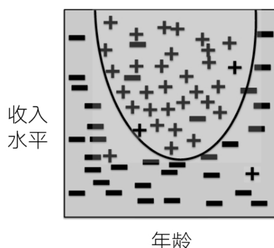
分类模型应用算法根据人们的年龄、收入、教育水平或在互联网上花费的时间等性质，将人划分为不同的类别。不同的算法意味着不同的属性与结果之间关系的基础模型。应用多种算法，也就是使用许多模型，能够产生更好的分类。

线性分类：在图a中，“正”（＋）代表参加投票的人，“负”（-）代表没有参加投票的人。在此基础上，可以用一个反映人们年龄与教育水平的线性函数来对某个人是否会参加投票进行分类。数据表明，受过更多教育的人更有可能去投票，同时年龄大的人也更有可能去投票。在这个例子中，图中的直线近乎完美地实现了分类。[7]



图a　用线性模型对投票行为进行分类

非线性分类：在图b中，“正”（＋）代表航空公司的常客（每年飞行超过1万英里的旅客），“负”（-）代表航空公司的所有其他旅客。中年人和收入更高的人更有可能乘坐飞机旅行。要对这些数据进行分类，需要先利用某个深度学习算法（如神经网络算法）找到一个非线性模型。神经网络模型包含多个变量，因此它们几乎可以拟合任何曲线。



图b　用非线性模型对航空公司的常客进行分类

决策树森林：在图c中，“正”（＋）表示参加科幻大会的人，基于他们的年龄和每个星期花在互联网上的小时数。在这里，我们使用了三棵决策树对数据进行分类。决策树根据各种属性不同的条件组合进行分类。图中显示的三棵决策树分别为：

决策树1：如果年龄＜30岁

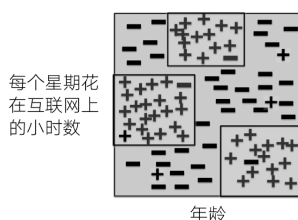
且每个星期花在互联网上的小时数介于［15，25］岁之间

决策树2：如果年龄介于［20，45］岁之间

且每个星期花在互联网上的小时数＞30

决策树3：如果年龄＞40岁

且每个星期花在互联网上的小时数＜20



图c　用决策树森林对会议参与者进行分类

树木的集合称为森林。机器学习算法会在一个训练集上随机构造出树，然后将那些在检验集上准确分类的树保存下来。

08　非线性模型

讨论非线性科学，就类似于讨论无大象的动物学。

约翰·冯·诺伊曼

在本章中，我们介绍非线性模型和非线性函数。非线性函数可以向下或向上弯曲，可以形成S形，还可以扭结、跳跃和波动。在下文中，我们将会讨论到所有这些可能性。现在先从依赖于凸性和凹性的模型开始讨论。我们阐明了增长和正反馈是如何产生凸性的，收益递减和负反馈又是如何产生凹性的。在绝大多数学科中，都包含了这两类模型。

关于生产的经济学模型假设交货期和库存成本会随着企业规模的增大而减少，从而使每单位产品的销售利润成了企业规模的一个凸函数，这也就解释了为什么沃尔玛能够获得如此高的利润。[1]关于消费的经济学模型则假设效用（或价值）是凹的，也就是说，第5块比萨带给我们的享受比第1块比萨小。在一个生态系统中，当一个新物种入侵并无须面对任何天敌时，其“人口”会以恒定的速率增长，这就产生了一个凸函数。但是随着“人口”的增长，它们的食物就会减少。因此，作为种群规模函数的适合度（fitness）是凹的。

本章由三个部分组成。第一部分讨论凸函数，包括了人口增长和衰退模型。第二部分讨论凹函数。在这一部分中，我们将会看到凹性意味着风险规避和对多样性的偏好。在第三部分中，我们研究了一系列经济学中的增长模型，它们结合了凹函数和线性函数。

凸函数

凸函数的斜率是递增的：函数值随度量值的增加而增加。例如，在一个人群中，可能结成的“对”的数量是这个群体人数的凸函数。一组3人，可以结成3个不同的“对”；一组4人，可以结成6个不同的“对”；一组5人，则可以结成10个不同的“对”。群体规模每增大一些，都会导致“对”的数量有更大的增加。与此类似，每一次，当厨师增加了一种新的香料时，他可以使用的香料组合数量就会增加很多。

我们要讨论的第一个凸函数模型是指数增长模型（exponential

growth model），它描述的是一个变量的数量（通常是指人口或资源）与它的初始值、增长率和周期数之间的函数关系。

指数增长模型

时间t的资源值V_t，其初始值为V_0，且以速率R增长，可以写成如下方程：

$$V_t=V_0(1+R)^t$$

这个单方程模型在金融、经济、人口、生态以及技术等领域中都发挥着核心作用。当我们把它应用于金融问题时，这里的变量就是货币。利用这个方程，我们可以计算出，年利率为5%时1 000美元债券在一年后的价值会增加50美元，而到第20年将增加100多美元。为了得出清晰的推论，我们假设增长率固定不变。根据这个假设，可以利用指数增长方程推导出72法则（Rule of 72）。

72法则

如果一个变量在每个周期内以R（增长率小于15%）的百分比增长，那么下面提供了一个很好的近似：

$$\text{翻倍所需的周期数}\approx\frac{72}{R}$$

72法则量化了最高增长率的累积效应。1966年，津巴布韦的人均国内生产总值为2 000美元，是博茨瓦纳的两倍。但是在接下来的36年里，津巴布韦几乎没有增长，而博茨瓦纳的年平均增长率则达到了6%，这意味着博茨瓦纳的人均国内生产总值每12年就会翻一番。于是在36年中，博茨瓦纳的人均国内生产总值翻了三番（增加了8倍）。因此，到了2004年，博茨瓦纳的人均国内生产总值达到了8 000美元，相当于津巴布韦的4倍。

同样是这个公式，还揭示了为什么房地产泡沫必定会结束而技术进步则不会。2002年，美国的房价上涨了10%。这个增长率意味着每7年翻一番。如果这种趋势一直持续35年，那么美国的房价将会翻五番，即增长32倍。这也就是说，一栋在2002年价格为20万美元的房屋在2037年将上涨到640万美元。当然，价格不可能一直这么涨下去，泡沫必定会破灭。与此不同，摩尔定律（Moore's law）则指出，可以安装在一块

集成电路上的晶体管数量每两年会增加一倍。摩尔定律之所以持续存在，是因为用于研发的投入带来了近乎恒定不变的进步速度。

人口学家则用指数增长模型研究人口问题。如果每年增长6%，那么人口在12年内就会翻一番，在36年内会翻三番，在100年内则会翻八番，即增长256倍。早在1798年，英国政治经济学家、人口学家托马斯·马尔萨斯（Thomas Malthus）就观察到了人口数量呈指数增长的现象，并在给出的一个模型中指出，如果经济体生产粮食的能力是呈线性增长的，就会出现粮食危机。短期的变化如下：人口增长模式为1、2、4、8、16、32……；而粮食生产的增长模式则为：1、2、3、4、5……。马尔萨斯预测，灾难很快就会发生。幸运的是，出生率不久之后就下降了，工业革命的到来也极大地提高了生产率。如果这两件事情都没有发生，那么马尔萨斯的预测应该是正确的。关键是，马尔萨斯忽视了创新的潜力。本章下面将要给出的模型重点就是创新，它颠覆了马尔萨斯担心的趋势。

指数增长模型也可以用于研究物种的增长，当然不仅仅适用于兔子。当你受到细菌感染时，那些肉眼不可见的细菌会以极高的速度繁殖。人类鼻窦中的细菌每分钟都在以4%的速度增加。应用72法则，我们可以计算它们每20分钟就会翻一番。一天之内，每个初始细菌都会繁衍出超过10亿的后代。[2]当然，由于鼻窦的物理空间有限，它们不可能一直繁殖，没有空间时增长就会停止。食物的限制、天敌的存在、生存空间的缺乏，都会减缓增长。有些物种，例如生活在美国郊区的鹿，或者被毒枭巴勃罗·埃斯科巴（Pablo Escobar）带入哥伦比亚的河马，虽然繁殖速度远远不如细菌，但是由于受到的限制很少，它们的种群迅速增大。[3]

具有正斜率的凸函数会以递增的值增加，具有负斜率的凸函数就会变得不那么陡峭，也就是说，最初具有较大负斜率的凸函数将逐渐走平。半衰期模型（half-life model）中的方程就是如此，这个模型可以用来刻画分解、折旧和遗忘。

在半衰期模型中，每H周期，数量就会衰减一半。因此，我们把H称为该过程的半衰期。对于某些物理过程，半衰期是恒定的。所有有机物都包含两种形式的碳：不稳定的同位素碳-14，以及稳定的同位素碳-12。在活的有机物中，这些同位素是以固定比例存在的。当有机体死亡后，体内的碳-14开始分解，其半衰期为5 734年，碳-12的数量则不会改变。美国物理化学家威拉德·利比（Willard Libby）意识到，通过测量

碳-14与碳-12的比例，就可以估计化石或人工制品的“年龄”，这种技术被称为放射性碳年代测定法（radiocarbon dating）。现在，古生物学家已经将放射性碳年代测定法应用于测定恐龙、猛犸象和史前鱼类遗骸的年代了。考古学家则用这种方法来判断古生物的真伪。利用这种方法，考古学家估计，在意大利阿尔卑斯山发现的冰人“奥茨”（Ötzi）的遗骸有5 000年的历史。而于1357年首次出现在公众眼前的“都灵裹尸布”则被认定为是14世纪的物品，而不是某些人所声称的用于耶稣基督葬礼的那一块。

半衰期模型

如果每H周期，剩余数量的一半会衰减，那么在t周期后，剩余的比例为：

$$\text{剩余比例} \approx \left(\frac{1}{2}\right)^{\frac{t}{H}}$$

半衰期模型的一个新应用是在心理学中。早期的心理学研究表明，人们几乎以接近固定不变的速度忘记信息。人们记忆的半衰期取决于事件的显著性。[4]2016年，电影《聚焦》（*Spotlight*）获得了奥斯卡最佳影片奖。假设，人们对奥斯卡获奖记忆的半衰期为两年，那么到了2018年，有1/4的人会记住这一事实；但是到了2026年，将只有1/1024的人还会记得这件事情。但对任何特定事件的回忆因人而异，对《聚焦》的导演汤姆·麦卡锡（Tom McCarthy）来说，他可能永远不会忘记他是哪一年获得奥斯卡奖的。

凹函数

凹函数与凸函数相反。凹函数的斜率是递减的。具有正斜率的凹函数会呈现收益递减的特点：当我们拥有的东西越来越多的时候，每个额外东西所能带来的价值会越来越少。几乎所有商品的效用或价值都呈递减趋势。闲暇越多、金钱越多、冰激凌越多，甚至与爱人共度的时光越多，对我们的价值就越小。一个直观的证据源于如下事实：包括巧克力在内，对任何事物的消费越多，我们就会越不觉得享受，同时愿意为它付出的代价也就越少。[5]

收益递减可以解释很多现象，包括为什么异地恋往往能够带来很大

的幸福感。如果你每月只能与你的伴侣相聚几个小时，那么每多一分钟都是一个莫大的惊喜。而在一个月不间断的相处后，幸福曲线的斜率就会变平，从而额外增加的相聚时间就变得不那么重要了。[6]同样的逻辑也可以解释为什么房地产开发商喜欢邀请人们在周末免费去他们的海滨公寓。在短暂的周末，你无法在海滩上享受足够长的时间，你会很想把房子买下。相反，如果让你在海滩上连续待上十天半个月，你可能就会觉得无聊。

当我们假设了凹性时，也就隐含地假设了对多样性和风险规避的偏好向。要证明前者，只需要给出一个有多个参数的凹函数就可以。如果人们的幸福曲线是凹性的，而且闲暇和金钱都在增加，那么人们就会更偏好休闲和金钱的组合，而不怎么喜欢只有金钱、没有闲暇或只有闲暇、没有金钱。而风险规避则意味着更偏好确定的有把握的事情而不怎么喜欢彩票，也就是不确定的事情。例如，一个厌恶风险的人会更喜欢得到100美元，而不是只有一半机会得到200美元，另一半机会什么也得不到。一个厌恶风险的人也更喜欢双层冰激凌甜筒，而不怎么喜欢要么没有冰激凌、要么可以得到四层冰激凌。

图8-1说明了为什么凹性就意味着风险规避。这幅图描绘了3种结果的幸福价值（幸福感）：高结果（H）、低结果（L），以及前两种结果的平均值（M）。给定形状向下的曲线，平均结果的幸福感会超过低结果和高结果的平均幸福感。凸函数情况下则相反。凸性意味着风险爱好：我们更喜欢的是极端值，而不是平均值。可以购买的股票数量是其价格的一个凸函数，因此，股票买家更喜欢价格波动。如果价格不断上涨和下跌，买家最终能够获得的股票比价格保持不变时获得的更多。[7]

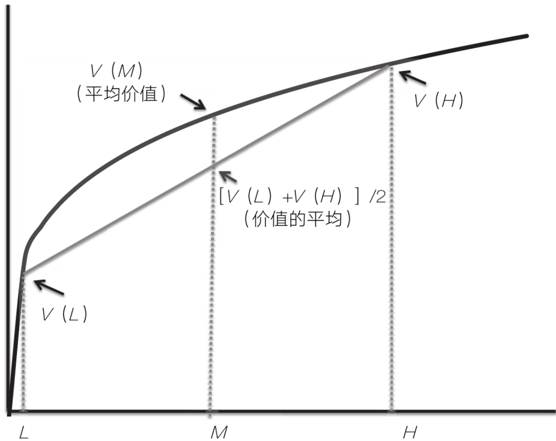


图8-1　风险规避：平均价值＞价值的平均

经济增长模型

接下来，我们将构建一系列经济增长模型。这些模型不仅揭示了增长的原因，还可以解释和预测各国的增长模式，还可以指导例如提高储蓄率等的行为。为了给对增长模型的研究奠定基础，我们先引入一个标准的经济生产模型，其中产出取决于劳动和实物资本。经验证据和逻辑都支持产出是劳动力和资本凹函数的假设。保持固定资本，随着投入的劳动力的增加，劳动力的价值应该变得越来越低。同样，在保持工人数量不变的情况下，添加更多的机器或计算机会增加更少的价值。逻辑推理还表明，产出应该是线性的，工人数量和资本总额翻番应该能使产出翻番。这就是说，一家拥有60名工人和一栋厂房的扫帚生产企业，在新建了一栋同样大小的厂房并多雇用了60名工人后，它的产出肯定应该翻番。柯布-道格拉斯模型（Cobb-Douglass model）是经济学中使用最广泛的模型之一，它同时包括了这两种性质。产出是劳动力和资本的凹函数，而且从规模上看是线性的。这个模型既可以应用于单个企业，也可以应用于行业或整个经济生产。[8]

柯布-道格拉斯模型

给定L个工人和K个单位资本，总产出如下所示：

$$\text{产出}=\text{常数}\times L^{a}K^{(1-a)}$$

其中a是介于0到1之间的实数，表示劳动力的相对重要性。

接下来，我们利用柯布-道格拉斯模型来构建经济增长模型。简化起见，我们假设经济体中有10 000名工人，并暂且不考虑工资和价格，这使我们能够专注于分析机器数量的变化是如何影响总产出的。然后，将资本投资与增长联系起来。为了使模型尽可能简单，再假设只生产一种商品——椰子。椰子含有丰富的椰汁和椰奶，可以作为食物。然而，椰子长在高高的树上，所以工人需要使用某种机器才能把它们摘下来。接下来，再做出一个非常不现实的假设，即机器本身也是用椰子制成的。这样可以简化模型，同时也保持了当前消费与未来投资之间的关键权衡。作为柯布-道格拉斯模型的一个特例，我们将生产函数写为工人数量的平方根乘以机器数量的平方根，即：

$$\text{产出}=\sqrt{\text{工人数量}}\sqrt{\text{机器数量}}=100\sqrt{\text{机器数量}}$$

如果经济体中只有一台机器，那么产量等于100吨。如果人们消费掉了所有100吨椰子，就不能投资制造新机器了，从而明年的产量将保持不变，也就是经济没有增长。如果他们投资1吨椰子制造了第2台机器，产量将增加到141吨，增长率为41%。如果他们制造了第3台机器，那么产量将增加到173吨。[9]通过不断地投资，经济增长率却逐渐下降，因此产出是一个凹函数。

我们已经大体知道投资是如何推动增长的。现在可以构建一个包含投资规则的更加精细的模型。假设投资等于储蓄率乘以产出，并假设机器按某个不变的折旧率折旧。例如，到了年底，不能再用的机器数量等于机器总数的某个固定比例。然后就可以得出，下一年的机器数量等于上一年的机器数量加上新投资的机器数量，再减少因折旧而减少的机器数量。于是，这个完整的简单增长模型由4个方程组成。

简单增长模型

产出函数：$O(t)=100\sqrt{M(t)}$

投资规则：$I(t)=s\times O(t)$

消费-投资方程：$O(t)=C(t)+I(t)$

投资-折旧方程：$M(t+1)=M(t)+I(t)-d\times M(t)$

其中，$O(t)$＝产出，$M(t)$＝机器，$I(t)$＝投资，$C(t)$＝消费，s＝储蓄率，d＝折旧率。

假设这个经济体中有100台机器，储蓄率为20%，折旧率为10%，产量等于1 000吨椰子，消费量等于800吨椰子，新投资200台机器。再假设，因折旧而损失的机器为10台，也就是在新的一年开始时将有290台机器。通过类似的计算可知，在第2年，产出将增长为1 702吨，而第3年的产出则将为近2 500吨。[10]由此可见，在这前3年，产出以递增的速度在增长。但是这种凸性只会在前几年出现，原因是机器的初始数量很少，因而折旧几乎完全不会产生任何影响。

随着时间的推移，机器数量的增加和折旧开始变得十分重要。从长远来看，产出的增长将完全停止（图8-2）。只要分析一下模型，就可以找到原因。投资是线性的，因为增加的新机器的数量是随产出呈线性增加的，同时产出则是机器数量的凹函数。因此，随着经济的增长，投资与机器数量的关系也是凹性的。然而，折旧与机器数量之间却是线性关

系。最终线性折旧会赶上产出的凹性增长。

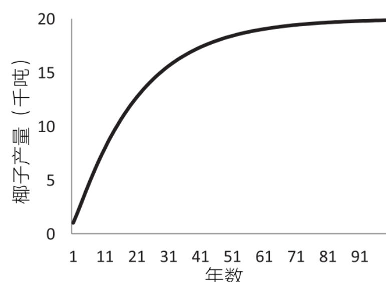


图**8-2**　简单增长模型中前**100**年的产出

在经济的长期均衡中，投资的新机器数量等于折旧损失的机器数量。在这个简单增长模型中，当经济体拥有40 000台机器并生产20 000吨椰子时，这种长期均衡就会出现。在这一点上，经济体在新机器上投入了20%的产出或4　000吨椰子，恰恰等于因折旧而损失的机器数量，也就是40　000台机器当中的10%。因此，因折旧而损失的机器数量等于通过投资和停止增长所创造的新机器数量。[11]

索洛*增长模型

现在构建一个更一般的模型，它是索洛增长模型（Solow　Growth Model）的简化，因此我们在索洛后加了一个星号。我们用实物资本取代机器，并将劳动力视为一个变量。此外，还添加了一个技术参数，它可以线性地增加产出。创新会使这个参数增大。与简单增长模型一样，当投资等于折旧时，长期均衡就会出现。不过，在这里，人们认为均衡时的产出水平取决于劳动力数量和技术参数，以及储蓄率和折旧率。[12]

索洛*增长模型

经济体中的总产出由以下方程给出：

$$产出 = A\sqrt{L}\sqrt{K}$$

其中，L表示劳动量，K表示实物资本量，A表示技术水平。长期均衡产出O^*由下面的方程给出：[13]

$$O^*=A(t)^2L\frac{s}{d}$$

长期均衡产出随劳动力数量的增加、技术的进步、储蓄率的提高而

增加，同时随折旧率的上升而下降。这些结果都不足为奇。更多的工人、更先进的技术和更多的储蓄理应可以增加产出，更快的折旧理应减少产出。但是，产量随着劳动力数量的增加和储蓄率的提高而线性增长这个事实却着实令人惊讶。劳动只能带来递减的收益，因此如果不考虑模型，我们可能会预测长期产出与劳动力数量之间是凹性关系。但是，随着劳动力数量的增加，产出也会增加，投资也会增加，进而带来更高的产出。从长远来看，投资的正反馈恰好抵消了收益递减。均衡产出与折旧率之间的关系是凸性的，折旧率降低20%可以使产出增加25%。

长期均衡产出也会随技术改进的平方而增加。因此，创新增加的产出要比线性增长更快。我们可以使用这个模型来了解其原因。如果从一个长期均衡的经济体开始，并将技术参数提高50%，那么产出将会增加50%，投资也将增加50%。然后投资超过了折旧，经济继续增长，投资将继续超过折旧，直到经济再增长50%为止，在这一点上，因折旧而造成的资本损失抵消了投资。这个计算过程揭示了创新乘数（innovation multiplier）的存在，创新有两个效应。首先，创新直接增加产出；其次，创新间接导致更多的资本投资，从而导致产出再次增加。因此，创新是持续增长的关键。[14]

需要注意的是，产出的这些增加不是瞬间发生的。当技术出现了一个突破时，技术参数的变化是相当缓慢的。直接效应的影响需要随着时间的推移显现。旧的实物资本必须被新技术的新实物资本所取代。当计算机技术进步刚刚发生时，一般企业的计算机不会马上变得更快，只有当技术发生了变化并且企业购买了新计算机之后，它们才会变得更快。实物资本投资增加导致的二阶增长则会发生在更长的时间范围内。技术与技术对增长的影响之间的滞后，可能意味着创新在出现后的几十年时间内都会导致增长。火车是在19世纪早期发明的，但是镀金时代（Gilded Age）并没有马上开始，直到19世纪的后半期才到来，这是一个长达50多年的“时滞”。另一个例子是，在阿帕网（ARPANET）出现整整30年后，互联网才开始步入繁盛期。[15]

国家缘何成功与失败

我们还可以将增长模型应用于一系列重大政策问题，例如，落后国家是否可以赶上发达国家？为什么有些国家取得了成功而有些国家却失败？政府在促进增长方面能够发挥什么作用？这些研究揭示了增长模型的价值和局限性。为了便于说明，不妨从低国内生产总值的国家实现快

速增长的能力开始讨论。模型证明，资本积累可以实现快速增长，技术投资也可以。一个实物资本较少的落后国家，有可能通过新的资本投入进入技术前沿，从而实现难以置信的高速增长。[16]

创新对长期增长来说是必不可少的，这种必要性也意味着一次性进口新技术有很大的局限性，而持续增长需要创新。

这些模型也表明，攫取和腐败，也就是政府将经济体的产出挪用于政府开支，会减少储蓄，进而削弱增长。对经济增长率的跨国比较研究的结果支持以下这些观点：减少攫取和腐败以及促进创新，都能推进经济增长。实现这些目标，需要一个强大但有限的中央政府来促进多元化。强大的中央政府能够保护产权、贯彻法治。多元主义能够阻止精英的俘虏，精英往往更喜欢现状，可能不会接受创新，因为创新往往可能具有很大的破坏性。

举一个破坏性创新的例子：克雷格列表网（Craigslist）。用户可以自行在这个网站上发布待售和求助广告。在21世纪初，克雷格列表网导致美国平面媒体行业失去数十万个工作岗位，但其实在那个时候，克雷格列表网本身只雇用了几十名员工而已。虽然许多人失去了工作，但是克雷格列表网通过增加技术参数使整体经济更有效率。而在一个多元化程度较低的社会中，平面媒体行业可能会游说政府禁止克雷格列表这样的网站。很显然，这样做将会减缓经济增长。

中国的经济优势

线性模型+72法则：从1960年到1970年，日本的国内生产总值以每年10%的速度增长。根据线性模型的预测，连年10%的增长，会使日本经济每7年翻一番（运用72法则）。1970年，日本人均国内生产总值约为2 000美元（以当前美元计）。如果这种增长趋势持续下去，那么到2012年，日本的人均国内生产总值将会翻六番，也就是说，人均国内生产总值将达到128 000美元。

增长模型：增长模型对日本经济增长的解释为实物资本投资的结果。这个模型还预测日本经济增长在一段时间内将会是凹性的。具体来说，这个增长模型的预测是，当日本的国内生产总值接近美国和欧洲时，日本的经济增长率会降低到1%～2%的跨国平均值。[17]证据支持这个预测。从1970年到1990年，日本国内生产总值的年增长率大约为4%。但是从1990年到2017年，它的增长率仅为1%或

更低。

中国的增长：从1990年到2010年，中国国内生产总值的增长率接近10%。2016年，中国的人均国内生产总值达到了8 000美元左右；正如增长模型预测的那样，增长速度已经放缓了。从2013年至2017年，增长率接近6%。同样，在中国，10%的增长率不可能一直持续下去，这与72法则相悖。如果中国经济在整个21世纪一直保持10%的增长水平，那么到这个世纪结束时，中国的人均国内生产总值将会超过1亿美元。

这毕竟是一个非线性的世界

之所以要构建非线性模型，是因为我们感兴趣的现象很少是线性的。在本章中，我们看到收益递减和收益递增是许多经济、物理、生物和社会现象的共同特征。我们还看到，在模型中包含曲率是有重要含义的。也许最重要的是，我们看到了函数形式能够影响我们的思维，用函数拟合数据有助于做出精确的表述。科学家可以使用碳-14数据来计算人工制品的年龄，经济学家还可以估计经济小幅增长的长期影响。

本章的一个核心结论是，一旦包括了非线性，直觉就变得不够用了。直觉可以告诉我们影响的方向：储蓄的增加、劳动力的增加和技术创新可以加快增长。模型还揭示了这些影响的形状和形式。正如我们所料，储蓄具有线性效应。从长远来看，劳动力的增加也是如此，即便模型假设短期收益递减。创新的增加还会产生乘数效应，我们对这种效应取其平方。第一个增长是创新的直接影响，产出的第二次增长则来自资本的增加。

在模型的帮助下，这些见解会变得很清晰。如果没有模型，我们通常可以推断出上升和下降的内容，但缺乏对功能关系形式的理解。我们会倾向于以线性的方式来思考，从而得出日本的经济将很快成为世界霸主的结论。利用模型，我们可以更好地思考非线性效应。也就是说，本章中介绍的凹函数型和凸函数，在非线性模型的巨大海洋中，只不过是沧海一粟。如果我们希望提高在复杂世界中推理、解释和行动的能力，就需要更深入地研究非线性现象。

09　与价值和权力有关的模型

你的价值不在于你知道了什么，而在于你能够分享什么。

罗睿兰（Ginni Rometty）

在本章中，我们讨论对个体行为者的价值和权力进行量化分析的模型。有些情况很容易处理。当一个群体的总产出等于每个成员个人贡献的总和时，每个人的价值就等于自己的贡献。但是，当集体产出不能分解为单独的组成部分时，例如当一组计算机程序员编写软件程序时，或者当一群创业企业家提出了新技术的某种创造性用途时，要分清每个人的贡献就会很困难。在美国，将权力授予政党时也会出现类似的问题：一方政党控制的议座数量与权力相关，但是这种相关并不是完美的相关。

在本章中，我们定义了度量价值和权力的两个标准。第一个标准是“最后上车者价值”（last-on-the-bus value，简称LOTB），它等于一位行动者在团队已经形成的情况下加入团队时的边际贡献。第二个标准是夏普利值（Shapley value），它等于行动者遍历所有可能的加入团队的序列，加入团队时的边际贡献平均值。例如，在一个由三个人组成的团队中，要求出一位行动者的夏普利值，先要求出他以第一、第二、第三位加入者的身份加入时的边际贡献，再计算平均值。我们是在合作博弈模型的框架下定义这些度量标准的。合作博弈模型由一组博弈参与者和一个价值函数组成。这个价值函数为每个可能的博弈参与者子集分配一个集体收益。

本章由四个部分组成。在第一部分中，我们定义了合作博弈、“最后上车者价值”和夏普利值，并给出了一些实例。在第二部分中，我们讲述了夏普利值的公理基础，并证明它是唯一能满足四个条件的度量标准。其中有两个条件分别是：对永远不能为团体增加价值的博弈参与者必须赋予零值，所有博弈参与者的价值总和必定等于博弈的总价值。在第三部分中，我们将夏普利值概念应用于执行某个创造性任务的团队。在这里，创造性的团队指每个成员都有新想法的团队。我们将阐明在这种情况下，夏普利值是怎样产生直观的价值衡量标准的。在第四部分中，我们考虑如何将夏普利值方法应用于投票博弈这个特殊情况。我们利用这个概念区分了投票权与投票百分比，结果发现两者之间并不总是一致

的。某个拥有20%席位的政党，在这一次投票中可能完全没有权力，但是在另一次投票中却可能得到1/3的总权力。

合作博弈

合作博弈由一组博弈参与者和一个价值函数组成。这个价值函数为博弈参与者的每个可能的子集（通常称为联盟）分配一个值。合作博弈的目标是刻画集体工作和联合项目。在合作博弈模型中，假设人们都会参与，以便我们可以专注于讨论如何为他们的参与分配价值。

合作博弈

合作博弈由N个博弈参与者和一个价值函数组成。这个价值函数为任何子集$S⊆N$分配一个值V（S）赋值。这些子集称为联盟。没有博弈参与者组成的联盟的价值等于零，即$V(\emptyset)=0$。所有N个博弈参与者的价值V（N）等于博弈的总价值。

在合作博弈中，一个博弈参与者的“最后上车者价值”等于当他是最后一个加入团队的人时，他所能增加的价值。“最后上车者价值”刻画了边际博弈参与者的价值。如果雇用4个人来搬运一张桌子，假设搬运这张桌子产生的价值为10，并且要4个人一起动手才搬得动，那么每个人的“最后上车者价值”均为10。如果只需要三个人就可以搬动这张桌子，那么每个人的“最后上车者价值”均为零。这里需要注意的是，“最后上车者价值”不一定是博弈的总价值相加。特别是，如果价值函数表现出了规模收益递减的性质，那么“最后上车者价值”的总和将小于博弈的总价值；如果增加的价值表现出了规模收益递增的性质，那么“最后上车者价值”的总和将超过博弈的总价值。

一个博弈参与者的夏普利值，等于他在所有可能加入的联盟的次序下对联盟边际贡献的平均值。换句话说，我们要在想象中按顺序将博弈参与者加入联盟中并计算每个博弈参与者为每个序列增加的价值。例如，考虑一家同时在西班牙和法国运营的小公司，它至少需要一位会讲法语的人和一位会讲西班牙语的人开展日常业务。假设该公司有三名员工：一名会讲西班牙语的人、一名会讲法语的人和一名既会讲法语又会讲西班牙语的双语人士。

现在假设，这个合作博弈为任何一位能讲法语和西班牙语的人分配

了1　200美元的价值。如果该公司能够运营，这个金额就等于公司每日的收入。如果任何两名员工来上班了，那么第三名员工就不是必需的。因此，在这个例子中，每个博弈参与者的“最后上车者价值”为零。

为了计算只会讲法语的那个人的夏普利值，我们要考虑这三个人来上班的所有6种可能的次序。在这6种次序中，只有在一种情况下，也就是只会讲西班牙语的人第一个到，然后这个只会讲法语的人第二个到时，这个只会讲法语的人才增加了价值。因此，这个只会讲法语的人的夏普利值就等于1/6乘以1　200美元，即200美元。与此类似，只会讲西班牙语的那个人只有当他第二个到且只会讲法语的那个人第一个到时，才能增加价值，因此他的夏普利值也等于200美元。而在其他四个次序中，既会讲法语又会讲西班牙语的人第一个到或者第二个到都能增加价值，因此，他的夏普利值等于800美元。所有这三个人的夏普利值总和等于1 200美元，也就是这个博弈的总价值。

夏普利值

给定合作博弈{*N*，*V*}，夏普利值的定义如下：

*N*个博弈参与者加入联盟的次序有*N*！个，让*O*代表这所有*N*！个次序。对于*O*中的每一个次序，将博弈参与者*i*增加的价值定义为当博弈参与者*i*加入时价值函数发生的变化。博弈参与者*i*的夏普利值等于他在*O*中所有次序上增加价值的平均值。

在了解了上述基本概念的基础上，现在可以构建一个更加复杂的例子了。想象一下，在赛艇比赛中，一个团队通常由四名桨手和一名舵手（舵手通常个子较小，控制划桨节奏和方向）组成。现在想组建一支赛艇队，就需要找到六名赛艇运动员，也就是合作博弈中的博弈参与者：五名高大强壮的桨手和一名舵手。在参加比赛的时候，四名桨手和一名舵手上场的团队价值为10；或者由五名桨手上场，但由于重量过重，整个团队表现不佳，这样的团队价值为2。

为了计算出夏普利值，假设这些博弈参与者以各种可能的顺序加入。如果舵手以第一、第二、第三或第四位的次序加入，那么他不会增加任何价值；当他以第五位的次序加入时，他增加的价值为10，这种情况出现的概率为1/6；如果他以第六位的次序加入，那么他将取代一位桨手，所增加的价值为8。将所有这些情况平均，可以发现舵手的夏普利值等于3。

而对于任何一个桨手来说，当且仅当他以第五位的次序加入时，才能增加价值，这种情况发生的概率为1/6。如果舵手没有加入，那么以第五位加入的桨手所增加的价值为2。如果舵手已经加入，那么以第五位加入的桨手所增加的价值为10。由于舵手最后一位加入的机会是1/5，同时舵手在前四位加入的机会是4/5，因此可以求出每一个桨手的夏普利值为7/5。[1]从直观上就可以看出，舵手的价值应该超过单个桨手的价值，同时考虑到桨手可以在没有舵手的情况下参加比赛（尽管成绩会“很差”），因此舵手的价值应该比所有桨手的总价值低。有无数种方法都可以在满足上面这两个约束的前提下分配价值，夏普利值只是给出了其中一个特定的分配方案：给舵手分配3，给所有桨手分配7。

夏普利值的公理基础

我们现在讲述夏普利值唯一满足的公理，这也就解释了为什么要优先考虑夏普利值而不是其他。第一，我们是通过对所有可能次序中博弈参与者的边际贡献来计算夏普利值的，因此任何永远不能增加价值的博弈参与者的夏普利值都为零。第二，对任何两个相同的博弈参与者，即对每个联盟贡献相同的任两个博弈参与者，也必须分配给他们相同的夏普利值。第三，由于所有次序的价值总和等于博弈的总价值，所以夏普利值的总和也必定等于与博弈的总价值。这里需要注意的是，“最后上车者价值”虽然满足前两个性质，但是却不满足最后一个公理。

在这三个公理的基础上，还可以增加第四个公理——可加性。这个性质要求，如果合作博弈的价值函数可以分解为两个价值函数，并把每个分解出来的价值函数分配给一个不同的合作博弈，那么复合博弈中一个参与者的价值应该等于他在两个分博弈中的价值总和。很容易看出，夏普利值也满足这个性质。不过，这四个公理唯一刻画了夏普利值这一点其实并不太明显。

证明一种度量唯一满足一组公理，也就为这种度量奠定了坚实的逻辑基础。如果没有公理基础，或许也可以找到某种直观的度量，但是我们可以认为它是武断的、似是而非的。上述公理告诉我们，如果选择任何其他度量，就不得不至少放弃其中一个公理。当然，这并不意味着夏普利值是唯一合理的标准。罗依德·夏普利（Lloyd　Sharpley）这位伟大的经济学家和数学家，可能是先写出了这个标准，然后才构造了这些只有它唯一满足的公理的。当然，谁先谁后其实并不重要。即便这些公理

是用后向方法构造出来的，只要我们接受，就应该采用这种方法。衡量标准的适当性取决于公理的合理性。就这些公理而言，前三个公理是无可争议的。第四个公理（可加性）虽然看上去比较复杂，但也是合理的，如果没有这个公理，博弈参与者就会有很强的动机去合并或分割联盟。

夏普利值的公理基础

夏普利值唯一满足以下公理：

零性：如果博弈参与者为任何联盟增加的价值都等于零，那么该博弈参与者的价值等于零。

公平性/对称性：如果两个博弈参与者对任何联盟都具有相同的增加价值，那么这两个博弈参与者具有相同的价值。

完全分配性：博弈参与者价值的总和等于博弈的总价值 $V(N)$。

可加性：给定两个定义在相同博弈参与者集合之上的博弈，它们的价值函数分别为V和$\hat{V}$，那么在博弈（$V+\hat{V}$）中，一个博弈参与者的价值等于该博弈参与者在V和$\hat{V}$的价值的总和。

夏普利值的应用

现在，我们将夏普利值应用在基于替代用途测试（alternative uses test）的合作博弈中。在测试中，每个人都必须为一种常见的物品想出一些新的用途，比如砖块。这种测试的目的是根据人们想出来的用途或用途类别来衡量一个人的创造力。我们在计算夏普利值的过程中，发现了一个直观的评分规则。

想象一下有三个人参加了某替代用途测试，分别是阿伦、贝蒂和卡洛斯。测试要求他们想出区块链的替代用途，这是一种分布式记账技术。如图9-1所示，阿伦和卡洛斯分别提出了6个想法，每个人的创造力得分均为6；贝蒂则提出了7个想法，因而得到7分。他们这三个人组成的团队的总创造力得分为9，因为总共有9个不同的想法（不同人提出的想法，有些是重合的）。

为了计算夏普利值，可以写下这个团队能够形成的所有6种可能的排序，而且只有当某个人为团队提供了独特的想法时才“给分”，然后再

对所有6种情况求平均值。或者，在计算夏普利值的过程中，我们可能已经注意到了，某个人因某个想法而“得分”的概率等于1除以所有提出了这个想法的人的数量。任何一个提出了一个他人没有的独特想法的人都可以获得满分。图9-1用粗体字来表示这类想法，例如阿伦提出的用区块链进行艺术交易的想法。

如果两个人提出了同一个想法，那么每个人都有1/2的机会首先加入该团队。同样，如果所有三个人都想到了同一个想法，那么每个人都有1/3的机会首先加入。这就是说，在想到同一个想法的人之间平等地分配得分能够产生夏普利值。因此，它是分配满足4个公理的值的唯一方法。这些值表明，虽然阿伦不是提出最多想法的那个人，但是他却增加了最多的价值。[2]

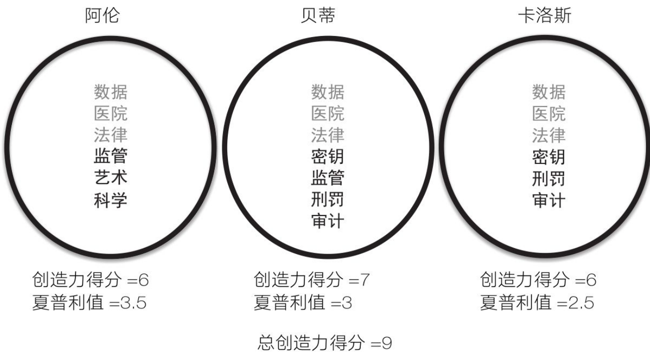


图9-1 在替代用途测试中应用夏普利值

夏普利-舒比克权力指数

接下来，我们将夏普利值应用于一类投票博弈。在这种投票博弈中，每个博弈参与者（代表某个政党或官员）控制着固定数量的席位或投票权，而且要采取行动，就必须获得多数席位或支持票。在投票博弈中，夏普利值通常被称为夏普利-舒比克权力指数（Shapley-Shubik index of power）。[3]通过对这个指数的计算，我们发现一个博弈参与者（政党）控制席位（投票权）的百分比与其权力之间并不存在直接的转换。

为了计算权力指数，考虑各个政党加入联盟所有可能的次序。如果某个政党加入了一个联盟并获得绝对多数，那么这个政党所增加的价值等于1。在这种情况下，我们就称这个政党是“关键的”。否则，这个政

党不会增加任何价值。

假设议会中共有101个席位，分别由4个政党掌握：A党控制了40个席位、B党控制了39个席位、C党和D党则各控制了11个席位。在这个例子中，如果A党首先或最后加入，那么A党就不会成为“关键的”政党。但是，如果A党在第二位或第三位加入，就肯定会成为“关键的”政党。因此，A党的权力指数为1/2。如果B党在第一位或最后一位加入，那么它就不能增加任何价值；如果B党在第二位加入，那么当且仅当A党已经在第一位加入时，B党才可能成为“关键的”政党。如果B党在第三位加入，那么它要想成为“关键的”政党，唯一的机会是A党在最后一位加入。这两个事件组合发生的概率分别为1/12。因此，B党的权力指数等于1/6。C党和D党也可以在两个与B党类似的事件组合中成为“关键的”政党。如果C党或D党在第一位加入，那么不可能成为“关键的”政党。如果A党在第一位加入，那么只要C党或D党在第二位加入，就能成为“关键的”政党。如果A党在最后一位加入，那么只要C党或D党在第三位加入，也能成为“关键的”政党。因此，C党和D党的权力指数也分别为1/6（图9-2）。

政党	席位数量	权力指数
A	40	1/2
B	39	1/6
C	11	1/6
D	11	1/6

图**9-2**　席位与权力之间的脱节

这个例子表明，一个政党控制的席位百分比与它实际拥有的权力之间可能存在着脱节。A党和B党控制的席位数量几乎相同，但是A党的权力却是B党的三倍；B党控制的席位虽然比C党或D党多得多，但是所拥有的权力却不比它们大。与这个例子相似的席位分配在现实世界的议会制度中经常出现。因此，只拥有少量席位的政党往往可以掌握很大的权力。例如，在以色列议会中共有120个席位。2014年，利库德集团领导的联盟共有43个席位，反对派联盟则拥有59个席位（仅略低于多数席位），正统派联盟拥有18个席位，但所有这三方都拥有相同的夏普利-舒比克权力指数。当然，相同的权力指数并不意味着正统派联盟在现实世界中确实拥有完全相同的权力，不要忘记，所有模型都是错的。但是它确实表明，正统派联盟的影响力超过了他们的席位数所占的百分比。

20世纪60年代中期，纽约拿骚县（Nassau County）监事会出现过惊人的席位与权力脱节的情况。该监事会由6名成员组成，每个成员控制的选票与该成员所代表的地区的人口成比例（图9-3）。投票事项要多数通过，需要得到115张票中的58张或以上。在三个最大的地区中，任何两个地区合作都可以稳占多数。因而另外三个地区的投票永远不可能是决定性的，这三个地区没有权力。例如，虽然北亨普斯特德（North Hempstead）地区拥有21票，超过了总票数的18%，但是并不能影响投票结果。

地区	票数	权力指数
亨普斯特德 1 区（Hempstead 1）	31	1/3
亨普斯特德 2 区（Hempstead 2）	31	1/3
奥伊斯特贝（Oyster Bay）	28	1/3
北亨普斯特德	21	0
长滩（Long Beach）	2	0
格伦科夫（Glen Cove）	2	0

图9-3　拿骚县席位与权力脱节

理论上说，夏普利-舒比克权力指数适用于任何席位或投票权分配不均等的情况，比如欧盟或美国的选举团，但这并不意味着它在所有情况下都是适当的方法。就美国选举团制度而言，50个州可以有50！（3×10^{64}）的不同次序。

当然，考虑到选民偏好的区域相关性，并非所有联盟都是可能的。密西西比州几乎不太可能与纽约州组成联盟。为了提供更有效的权力衡量标准，我们需要将某些联盟置于更优先于其他联盟的位置，或将某些联盟排除出去。在后面的章节中，我们描述了允许排除某些联盟的迈尔森值（Myerson Value）。

小结

个体的夏普利值与为联盟增加的平均贡献相对应。它是衡量增加价值的一种标准。在投票博弈中，也可以将夏普利值解释为权力的一种度量。不过，夏普利值可能并不一定总是最好的衡量标准。假设威胁是可信的，那么在一个群体已经形成的情况下，个人的“最后上车者价值”可能是衡量权力的一个更好标准，因为它能够度量每个人通过威胁离开可

以攫取多大利益。

在这些情况下，联盟会希望减少“最后上车者价值”。通过扩大联盟规模，可以创建出一个具有很高的总价值、同时“最后上车者价值”又足够低的联盟。不断加入新成员，会使现有成员变成“可以放弃的”，从而使“最后上车者价值”趋向于零。我们在实践中确实可以观察到这一点。例如，雇主会通过雇用多余的工人来削弱工人的权力，制造业企业会向多个相互竞争的供应商采购中间产品，政府会与多个承包商签订合同，等等。

同样的直觉也可以用于解释美国立法机构中出现的联盟。国会游说者和政党领导人希望通过法案（价值的一种结果），同时又试图限制个别众议员和参议员的权力。[4]如果游说者努力争取到了通过法案所必需的最低数量的众议员和参议员的支持，那么每一个众议员和参议员都会拥有很大的“最后上车者价值”。任何一个人都可以通过改变自己的投票来推翻那个法案。在这种情况下，游说者可以通过收买绝大多数众议员和参议员来降低他们的“最后上车者价值”。同样的逻辑也意味着，只拥有微弱多数的政党可能是非常难以驾驭的，因为每一个成员都拥有很大的“最后上车者价值”。而在某个政党拥有了绝大多数席位（投票权）的时候，没有任何众议员或参议员能够拥有太大的权力。

将视野放大到现代互联网世界，我们发现应用“最后上车者价值”和夏普利值的概念来思考权力问题非常有用。无论是个人、组织、企业，还是政府，抑或是恐怖组织的权力，都部分取决于偏离合作制度可以造成的损害的程度，也就是“最后上车者价值”。一个技术高超的计算机黑客，由于拥有摧毁大量财富的力量，因而拥有巨大的权力。即便黑客完全不能给社会创造价值，这个结论依然成立。

在考虑跨国企业或其他跨国组织的价值时，夏普利值可能是一个更好的衡量标准。在这些情况下，退出本身就是一个不可行的选择。能源公司必须参与能源生产博弈、能源分配博弈、房地产博弈、环境博弈、就业博弈等。这样的公司的总增加值等于各个领域的增加值之和。

通过合作博弈论的视角来思考权力和价值，可以得出很多深刻有力的洞见。合作博弈还指出了我们下一步应该关注的地方。在政界和商界，并不是所有联盟都是合理的。不过目前的模型假设它们合理。更丰富的模型需要考虑到世界的连通性。咨询公司和金融公司要从科技公司购买软件，科技公司和咨询公司通过金融公司进行投资和借贷，金融公司和科技公司要聘请咨询师。在这些网络中，每个参与者都能增加价值

并发挥影响力。要计算出这种环境下的权力，我们需要网络模型。

10 网络模型

网络理论是科学的一个完整分支。但是就过去的二三十年来说，它相对较新。我们还没有机会把所有这些理论从大学中拿出来，然后问自己："我们应该建立什么样的网络？应该将网络用于什么样的目的？"

安妮-玛丽·斯劳特（Anne-Marie Slaughter）

本章将介绍网络模型。对网络进行全面研究，需要写好多本书。因此，我们在这里只专注于一个更加温和的目标：只希望了解有关网络的基础知识，能够给网络的各个组成部分命名，并讨论它们对于建模的重要性。我们得出的答案是，网络几乎总是很重要。我们构建的任何模型，无论是市场模型、传染病传播模型，还是信息传播模型，都可以通过将参与者嵌入网络中而变得更加丰富。[1]

网络无处不在。人们经常会谈起贸易网络、恐怖主义网络，以及志愿者网络。不同物种会组织成食物链，那是一种网络形式。企业会建立供应链，那也是一种网络。如前所述，将金融系统视为一个支付承诺的网络会很有效。网络对于理解社会关系一直都很重要。在人类历史的大部分时间里，社交网络受到地理限制，难以扩展。由于技术的进步，许多社会互动和经济交易现在都是通过虚拟网络进行的，并且可以使用模型进行分析。

本章内容仍然遵循前面用来讨论分布时所用的模式，即结构—逻辑—功能。我们首先用一系列统计量来表征网络结构，包括：度、路径长度、聚类系数和社区结构等。然后我们讨论了一些常见的网络类别：随机网络、中心辐射网络、地理网络、小世界网络和幂律网络。之后探索网络形成的逻辑，我们构建了一些微观层面的流程，生成所能观察到的网络结构。最后讨论功能，也就是网络结构为什么是一个重要的问题。

本章主要关注网络结构的五个重要含义。我们首先从友谊悖论入手，在对它进行了全面分析之后，描述六度分隔理论和弱关系属性的强度。最后讨论了网络在节点或出现故障时的鲁棒性，解决了网络上的信息集结问题。本章最后还讨论了网络会如何影响模型结果。

网络的结构

网络由节点以及连接节点的边（edge）组成。由边连接起来的节点互为邻居。如果沿着边，可以从任何一个节点到达任何其他节点，就将这样的网络称为连接的网络。网络可以用图形来表示，也可以用边的列表表示，或者也可以用由0和1组成的矩阵表示，其中第A行、第B列的一个数字表示节点A和节点B之间的边。虽然人们更喜欢用图形来表示网络，但是其实用列表和矩阵来表示网络，才更适用于计算网络统计数据。

网络中的边可以是定向的，也就是说，可以从一个节点指向另一个节点。在信息网络中，一条有向边表示一个人从另一个人获取信息。在生态系统网络中，从红尾鹰到灰松鼠的一条有向边表示红尾鹰吃松鼠。边也可以是非定向的；连接两个朋友的边就是非定向的。在非定向网络中，一个节点的度（degree）等于连接到它的边的数量。

网络以一组网络统计数据为特征。对于每个统计量，我们可以计算网络平均值和所有节点的分布。例如，友谊网络的平均程度告诉我们平均每个人有多少个朋友。度分布（degree distribution）告诉我们某些节点是否比其他节点连接得更多。社交网络的分布比万维网、互联网和引文网络更加平等，后面这几类网络都有很长的尾巴。

路径长度，指两个节点之间的最小距离，与度成反比。当增加边时，就缩短了节点之间的平均距离。在航空公司的航线网络中，路径长度对应于人们从航线网络中的某个城市到另一个城市所需的航班数量。如果要在两家航空公司之间做出选择，在其他所有条件（即价格）都相同的情况下，旅客会更喜欢平均路径长度更低的那家航空公司。平均路径长度也与信息丢失相关。经过多人中转传递的信息比直接在两个人之间传递的信息更容易遭到扭曲。最短路径上的节点在网络中起着关键作用。如果信息是通过最短路径传递的，那么就必定会经过最短路径上的节点。节点的介数得分（betweenness score）等于通过该节点的最小路径的百分比。在社交网络中，介数得分高的人掌握更多信息并且拥有更多权力。

最后一个统计量是聚类系数，它等于节点的邻居节点对当中，同时彼此也互为邻居节点对所占的比例。例如，一个人有10个朋友，这些朋友可以组成45个对。如果在这45个对当中，有15个对本身也是朋友，那么这个人的聚类系数就等于1/3。如果所有这45对都是朋友，那么这个

人的聚类系数就等于1，这也是所有可能值当中最大的一个。整个网络的聚类系数等于各个节点聚类系数的平均值。

网络统计量

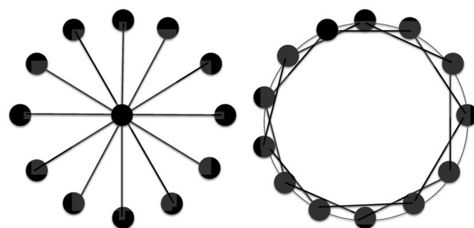
度：节点的邻居数（即边数）。

路径长度：从一个节点到另一个节点必须遍历的最小边数。

介数：经过某个节点连接两个其他节点的最短路径数量。

聚类系数：一个节点的邻居对当中，同样也由一条边连接的邻居对所占的百分比。

图10-1显示了一个辐射网络和一个地理网络，它们各具有13个节点。在这个辐射网络中，中心节点的度为12，所有其他节点的度均为1，因此平均度小于2。这种度分布是“不平等”的。中心节点与其他每个节点的距离均为1。所有其他节点与中心节点的距离为1，与中心节点之外的任何一个节点的距离为2。因此，这个辐射网络的平均路径长度也小于2。中心节点位于任何两个其他节点之间的最小路径上，于是介数得分为1。任何一个分支节点都不位于连接其他节点的任何最小路径上，因此它们的介数得分均为0。最后，在这个辐射网络中，连接到某个节点的任何节点都不彼此连接。因此，网络的聚类系数为0。



网络测度	辐射网络		地理网络
	中心节点	分支节点	
度	12	1	4
平均路径长度	1	$1\frac{11}{12}$	2
介数	1	0	$\frac{1}{12}$
聚类系数	0	0	$\frac{1}{2}$

图10-1　一个中心辐射型网络和一个地理网络

在这个地理网络中，每个节点都连接到位于它右侧和左侧的两个节

点，因此平均度等于4。每个节点到4个节点的距离为1、2、3，因此平均距离恰好等于2。从图10-1可见，这个地理网络的度和距离分布都是简并性（degenerate）的，因为每个节点都具有相同的度和相同的平均距离。可以看出，每个节点的介数都等于1/12。[2]每个节点都有4个邻居，可以构成6个对。在这6个对中，恰好有3对是相互连接的：直接靠着该节点的左右两个节点分别连接到再外一点的节点，并相互连接。因此，聚类系数等于1/2。

刻画网络的聚集程度的另一种方法是将节点划分为不同的社区（community）。在年轻人的友谊网络中，这里所说的社区可能对应于对艺术、体育或科学感兴趣的青少年。或者，社区也可以通过种族和性别来定义。政治联盟网络可能可以分为地区性的或意识形态上的盟友。可以用来确定社区的方法有很多，其中一种方法是依次移除具有最高介数的边，因为介数高的边更有可能将不同的聚类连接起来。还有一种方法是将社区的数量视为给定的，并在特定的目标函数下寻找最佳划分方法，例如最小化社区之间边的数量或最大化社区内部边的比例。[3]

我们可以使用社区检测算法（community detection algorithm）去查看网络数据的问题。有研究表明，人们可能会生活在“网络泡沫”（online bubbles）当中。这也就是说，我们所属的社区，可能是由只从类似来源获取新闻的人组成的。如果真的是这样，那么无疑会对社会凝聚力产生重要的影响。在互联网出现之前，情况可能也是这样的，但是很难用数据证明。现在，数据科学家可以在网络上抓取海量数据，从而将人们的新闻来源准确地识别出来，实际上我们确实在一定程度上生活在“泡沫”当中。现在，模型可以给出社区的正式定义，数据可以揭示这些社区的力量。结合我们的判断力，就可以根据数据做出明智的推断。

常见的网络结构

在分析网络时，我们遇到了网络过于多样性的问题（图10-2）。少数几个网络统计量无法确定具体的网络结构：人们可以构建出数十亿个具有10个节点且平均度为2的不同网络。还可以通过检验它的统计指标是否与某个常见的网络结构有显著差异来表征网络。例如，当一位法官引用了另一位法官的判决时，一个研究者可能会收集司法引用的数据并将其以网络的形式表示。这种网络从图形上看似乎有一个很特殊的结构和集群。我们可以将这个网络的统计数据与具有相同数量的节点和边的

随机网络进行比较，以检验这个网络是不是随机的。在第一种常见的网络随机网络（random network）中，聚类系数等于一条随机的边的概率，因为一个节点的两个邻居并不比任何其他随机选择的节点更可能包含一条边。

随机网络的蒙特·卡罗方法[(5)]

为了检验一个具有N个节点和E条边的网络是不是随机网络，可以创建大量具有N个节点和E条边的随机网络，并计算出度、路径长度、聚类系数和介数的分布。然后，执行标准的统计检验，以确定接受还是拒绝那个网络的统计数据可能抽取自该模拟分布的假设。[4]

理论模型通常假设某种特定的网络结构。有的研究者偏好假设随机网络，而有的研究者则偏好假设规则的地理网络，例如这样的网络：节点排列成圆形并且每个节点在每个方向上都连接到最近的节点。这也是第二种常见的网络，有一种地理网络将节点排列在棋盘上，并让每个节点与自己东、南、西、北的邻居相连。大多数常见的地理网络都具有较低的度，即节点仅连接到本地邻居，并且具有相对较大的平均路径长度。在地理网络上，介数和聚类系数不会有变化。

第三种常见的网络是幂律网络，这种网络的度分布是幂律的。少数节点有许多连接，同时大多数节点的连接则非常少。第四种常见的网络是小世界网络，它结合了地理网络和随机网络的特征。[5]要想构建一个小世界网络，可以从一个地理网络开始，然后进行“重新布线”，方法是随机地选择一条边并把这条边所连接的其中一个节点替换为一个随机的节点。如果“重新布线”的概率等于零，所拥有的就是一个地理网络；如果“重新布线”的概率等于1，那么就有了一个随机网络；而当概率介于这两者之间时，就会得到一个小世界网络，以小集群区别于通过随机链接连接到其他集群的地理网络。社交网络看起来类似于小世界，每个人都有一群朋友，以及若干随机的朋友。

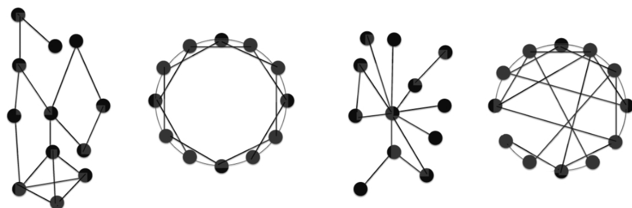


图10-2　随机网络、地理网络、幂律网络和小世界网络

网络形成的逻辑

现在简要讨论一下几个描述网络形成的模型，这些模型给出了解释网络结构的逻辑。我们遇到的大多数网络结构都是从个体行为者做出的关于建立连接的选择中涌现出来的。友谊网络、万维网和电网都是如此。这些网络不是计划的结果。不过，也有一些网络，例如供应链网络，确实是计划的产物。我们希望按计划构造的网络对节点的故障具有鲁棒性。当然，自发涌现的网络结构都具有鲁棒性这个事实是一个谜。

我们已经讨论了如何创建随机网络和小世界网络。只需要随机创建一组节点，然后绘制连接随机节点对的边，就可以创建随机网络。通过构建一个规则的地理网络（常用的方法是在一个圆周上排列节点并在每个方向上连接k个邻居），然后随机“重新布线”一部分边，就可以创建一个小世界网络。

电网形成的模型依赖于经济和工程原理。电网必须为家庭、企业和政府提供电力。无论生产者是营利性企业还是公共事业公司，都没有动力去创造聚集程度很高的网络，因为那种网络效率低下。但是，聚类的缺乏降低了网络的鲁棒性。经济和工程方面的考虑也排除了远距离跳跃——跨越网络连接的可能性。电力公司不会建立从芝加哥到达拉斯的直接连接。然而，人和企业却可以，一个芝加哥人可能会与达拉斯的某个人建立起友谊，新加坡的一家公司可能会与底特律的一家公司进行交易。正如后面将看到的，这些远距离跳跃有助于提高网络的鲁棒性。

要创建一个具有长尾分布的网络，可以利用优先连接模型的一个变体。先随机创建一些节点，然后画出从新节点到现有节点的边。如果我们令连接到节点的概率与节点的度成正比，就可以产生幂律的度分布。在这个模型中，越早“到达”的节点的度越大。但是这个模型有一个缺点，那就是，它不允许节点质量有任何差异。更高质量的节点本应具有更高的度。不过，质量和度的网络形成模型（quality and degree network formation model）纠正了这种缺憾，并且产成了长尾分布。

质量和度的网络形成模型

创建d个互不连接的节点。在每周期t中创建一个质量为Q_t的、从分布F中抽取出来的新节点。根据其他d个节点的度将这个新节点连接到那些节点上。用D_{it}表示在时间t时节点i的度，那么给定N个节点时选择节点i的概率等于：

$$\frac{D_{it}+Q_{it}}{\sum_{j=1}^{N}(D_{jt}+Q_{jt})}$$

如果新节点质量的均值和方差都足够低，那么这个模型就类似于标准的优先连接模型。如果质量分布有一条长尾，那么质量很高的新节点的度可以增长到非常大的程度。[6]

网络的功能

前文中，已经提到过友谊悖论；事实上，在任何网络上，平均来说，人们确实不可能比他们的朋友拥有更多的朋友。我们可以利用辐射网络来说明出现这种情况的原因及其背后的逻辑。在辐射网络中，12个人中的每人都只有1个朋友，只有1个人有12个朋友。有1个朋友的那12个人都连接到了中心节点，中心节点有12个朋友。这个特征，也就是度更高的人与更多的人连接在一起的事实，驱动了结果。在中心网络上，平均来说，所有人只有不到两个朋友。然而，平均而言，每个人的朋友都有超过11个朋友。

友谊悖论适用于任何网络：电子邮件网络、学术引文网络、银行网络和国际贸易网络等。平均而言，一篇学术论文引用的参考文献被引用的次数比这篇文章本身更多；与一个国家的贸易伙伴进行贸易的国家数量，要比与这个国家进行贸易的国家更多；食物网络中与单一物种相连接的多个物种的连接比该物种自身更多。在具有更加分散的度分布的网络上，朋友的数量与朋友的朋友的数量之间的差异会变得更加明显。例如，根据Facebook的数据对友谊网络进行的一项研究结果表明，一个人平均大约有200个朋友，而他们的朋友平均来说有超过600个朋友。[7]

友谊悖论

如果网络中任何两个节点的度不同，那么平均而言，节点的度会低于其相邻节点。换句话说，平均而言，人们的朋友比他们自己更受欢迎。[8]

友谊悖论的逻辑可以扩展到任何与朋友数量相关的性质。如果活跃、快乐、聪明、富有和友善的人平均而言会拥有更多的朋友，那么一个人的朋友平均来说会更活跃、更快乐、更聪明、更富有、更友善。[9]想象一下这样一个网络，在不快乐的人中，90%的人有4个朋友、10%的

人有10个朋友。而快乐的人的情况则相反：在快乐的人中，10%的人有4个朋友，90%的人有10个朋友。人们的朋友将不成比例地由拥有10个朋友的人组成。由于那些人中的大多数都是快乐的，因此绝大多数人的朋友都会比自己更快乐。

现在，我们来看一下六度分隔理论，也就是地球上的任何两个人都可以通过6个或更少的朋友联系到一起。虽然友谊悖论适用于任何网络，但是六度分隔却只适用于某些类型的网络。这个术语源于美国社会心理学家斯坦利·米尔格兰姆（Stanley Milgram）在20世纪60年代进行的一项实验。米尔格兰姆向内布拉斯加州奥马哈市和堪萨斯州威奇托市的296人寄出了一些包裹，那些包裹最终需要转寄给在马萨诸塞州波士顿市的一个人。收到包裹的人必须遵守相同的规则：所有参与者只能通过邮政系统将包裹寄给他们认识且他们认为更有可能认识那个波士顿人的人，并附上同样的指示。每个参与实验的人都要在一份记录路径的名册上签名，并邮寄明信片给研究者，以便研究者可以跟踪链条上的断点。最终，有64个包裹抵达了波士顿。这些抵达波士顿的包裹所经历的平均路径长度略小于6，因此就有了“六度分隔”这种说法。

50年后，研究者们又组织了一场更大规模的“六度分隔”实验。这一次是利用电子邮件进行的，实验组织者在全球范围内设定了18个最终收件人，邀请了20 000多个人参加了这个实验。电子邮件链的中间路径长度在5到7之间，具体取决于最初的发送者与最终目标之间的地理距离。由于实验中运用的路径长度不等于实验参与者之间的最小路径长度。因此，这些证据表明，大多数人之间用不着6度就可以连接起来了。[10]

在这里，我们构建了一个简化版的小世界网络，以便直观地理解六度分隔理论。这个小世界网络假设每个人都有一个由若干个圈内好友构成的小群体，这些人彼此认识，而且每个人都拥有不属于这些圈内的朋友，我们把这些圈子外的朋友称为“随机朋友”（random friends）。[11]图10-3表明，某人（用黑色圆圈表示）有5个圈内好友和两个随机朋友。它还显示了这个节点的朋友（用浅灰色圆圈表示）的部分“朋友圈”。

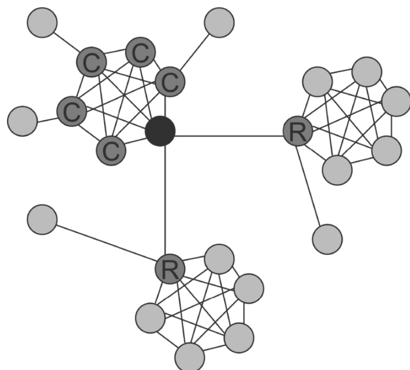


图10-3　一个节点的圈内好友（C）和随机朋友（R）

这些随机朋友也可以认为是一种弱关系，他们可以将你连接到其他群体的人。我们的弱关系，也就是网络中的随机朋友，由于连接了具有不同兴趣和信息的社区，从而发挥了重要的信息作用。因此，社会学家很强调弱关系的力量。[12]

在这种网络结构中，我们可以计算出二度邻居，也就是朋友的朋友的数量，方法是将随机朋友的所有朋友人数相加，但是不把圈内好友计算在内，因为他们本身就是节点群体的成员。与此类似，我们可以计算出朋友的朋友的朋友的数量，方法是把所有“圈内好友”的随机朋友的朋友加进来，但是不把随机朋友的“圈内好友”的“圈内好友”计算在内，因为他们在计算二度邻居时已经加进来了。为了产生六度分隔，我们将相同的逻辑应用于一个具有100个“圈内好友”和20个随机朋友的网络。

六度分隔

假设每个节点有100个“圈内好友”（C），他们彼此都是朋友；以及20个随机朋友（R），他们没有与节点共同的朋友。

一度：C＋R＝120

二度：CR＋RC＋RR＝2 000＋2 000＋400＝4 400

三度：CRC＋CRR＋RCR＋RRC＋RRR＝328 000

四度[13]：17 360 000

五度：＞10亿

六度：＞200亿

由于假设随机朋友的朋友之间没有重叠，这个模型隐含地假定人口

是无限的。但是在现实世界中，随着度数的增大，真实的社交网络会出现朋友之间的重叠。在包括了重叠和其他真实世界特征（例如朋友数量的异质性）的网络中，实际值将会与上面计算出来的值不同。不过，每个度的邻居数量的相对大小将保持相似：一个人的三度邻居（朋友的朋友的朋友）会比二度邻居（朋友的朋友）多得多。

三度朋友可能是相当重要的，在我们这个例子中，三度朋友的人数超过了25万。与一个人的“圈内好友”不同，一个人的三度朋友往往会住在不同的城市，就读于不同的学校，拥有不同的信息，他们会更加多样性。他们也足够接近，可以建立起信任关系：朋友的朋友的朋友可能是你的室友的母亲的同事，或者是你妹妹的男朋友的姨妈。三度朋友的数量很重要，他们的多样性以及相对接近性使他们成了你的重要资产，他们可以提供新的信息和工作机会。这些人最有可能帮助你找到工作，促使你搬到新的城市，或者成为生活中、商业上的伙伴。

网络结构的鲁棒性

现在来讨论网络结构的鲁棒性，看看当网络的一个节点或一条边发生故障时会怎样。网络最重要的性质是，它在受到冲击时是不是仍然能保持连接。我们可以使用模型来计算网络保持连接的概率——作为移除节点数量的函数。还可以考察当移除某些节点时平均路径长度会发生什么变化。例如，在航空公司的航线网络中，这种分析可以告诉我们，如果某个机场由于天气原因或电力故障而关闭了，将会需要多少额外的航班。

在这里，我们考虑网络中最大连接部分，即巨型组件的大小怎样随着节点的随机失败而变化的问题。图10-4显示了一个大型随机网络和一个大型小世界网络中巨型组件的大小。在这个随机网络中，巨型组件的大小一开始会呈线性下降。在边概率等于1除以节点数这个临界点上，巨型组件的大小会下降到原始网络的任意小的比例。但是，小世界网络却不会出现这种突然的变化。因为在小世界网络中，大多数连接都存在于地理聚类中。每个聚类都可以顶住多个节点发生故障的冲击。这个性质与随机连接相结合，可以有效地防止整个网络崩溃。

从图10-4可以看出：缺乏局部聚类的稀疏网络更容易出现故障。我们可以将这个结论应用于对电网的分析。电网缺乏远距离跳跃连接和紧密聚类，而保证小世界网络鲁棒性的就是这些。在电网中，节点或连接

如果发生了故障，是不能通过聚类中的其他连接或与仍然有效节点的远距离连接来克服的。因此，局部故障就级联放大地传播到整个网络。[14] 与此形成鲜明对照的是互联网。具有长尾度分布的互联网对随机节点故障具有很强的鲁棒性。互联网的度分布意味着，绝大多数节点的连接很少，因此即便它们发生了故障，网络也能保持连接。

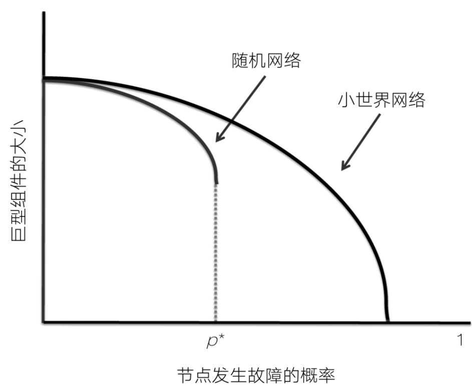


图10-4　巨型组件的大小作为节点故障的函数

到目前为止，我们一直假设节点是随机发生故障的。我们也可以考虑移除战略性节点。有些具有长尾分布的网络，比如说互联网，现在也变得比以前脆弱了。战略性地移除度最高的节点会破坏整个网络。可以从辐射网络看出其逻辑。当随机移除节点时，辐射网络仍然可以保持连接，除非直接移除中心节点，但那是一个低概率事件。然而，战略性地移除节点，即直接毁灭中心节点，却可以一步切断网络连接。

对于某些网络，例如恐怖主义网络和毒品贸易网络是我们想要切断的。如果这些网络是像电网那样的稀疏网络或者具有长尾度分布，那么就可以通过移除战略节点来断开这些网络。对于恐怖主义网络来说，这意味着必须逮捕拥有链接最多的恐怖组织成员。如果这些网络类似于小世界网络，那么它们就会具有鲁棒性，甚至在战略性的节点被移除之后仍然能屹立不倒。尝试切断这种网络任何“区段”的努力都将失败，因为随机重新连接的存在会将这些“区段”重新连接到网络上。

小结

当我们尝试构建关于由人构成的网络模型时，目的通常是用它们来刻画社会影响，也就是在社会网络中，一个人的成功、行为、信息或信

念，会影响他们的朋友的成功、行为、信息或信念。一个人的行为既可能是依赖于情境的，也可能是由内在因素决定的；个人对共同事业的价值或贡献也是如此。一个人的价值或贡献可能是源于他本身的某种性质，例如聪明才智、努力水平或好运气，但是，一个人的成功或许也可以归功于其朋友和同事的网络。这其实是一个非常古老的问题：成功到底取决于你所知道的东西，还是取决于你所认识的人？

试想一下，一群科学家在一个实验室里一起进行研究工作。他们分享想法和知识，互相提建议。因此，某个科学家发表的学术论文、申请的专利或取得科学突破的数量不仅取决于他自己知道些什么，而且也依赖于他与其他科学家的互动，他会受到他所认识的人的影响。我们要把环境特征和内在性质一起考虑，然后再来确定某位科学家的成功应该在何种程度上归因于自己的努力程度。

成功主要取决于人才的信念，有些投资公司雇用明星基金经理来操盘，但是往往无法取得预期的结果。经验证据表明，顶级投资者也依赖于向他们提供特定类型信息的同事网络。[15]放在更大的文献背景内，这个发现其实并不令人意外。许多文献（其中有一些是基于模型的）都表明，一个人在一个组织中的位置会影响他的成功。

当然，成功肯定与能力相关。一个能够让投资者获利数百万的商业创意当然是一个好主意，发表数百篇论文并获得无数奖项的科学家肯定具有很高的科研能力。但是另一方面，我们也不能否认，恰恰是那些在网络中占据了最核心位置的人做出了最大的贡献。我们可以使用介数和其他中心性测度来衡量一个人的位置。在网络中占据了介数很高位置的那些人，填补了著名社会学家罗纳德·伯特（Ronald Burt）所说的社区之间的“结构洞”（structural holes），我们可以使用特定的算法识别出这些结构洞。[16]由于能够从多个社区获取信息和思想，这些填补了结构洞的人拥有很大的权力和影响力。当然，要想去填补结构洞，你得有相当高的才华和能力才行。看到一个洞就跳下去并不算填补结构洞。要填补结构洞，你必须让社区的每一个人信任和理解你，你必须熟悉每个社区的知识库。

运用同样的原理，我们还可以对企业的价值进行评估并讨论国家的权力分配。我们可以将企业的价值视为内在价值，着重从资产负债的角度分析。还可以考察该企业运营的情境，例如它在供应链中的位置。与此类似，一个国家的权力取决于它的资源和联盟。无论是对于企业还是对于国家，它们的内在性质都是与连通性相关的。那些在网络中占据很

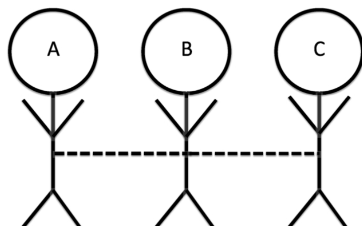
高位置的人也拥有某些举足轻重的特性。

这里的分析以及大多数文献都将节点视为分析的单位。事实上，边同样也很重要。而且，从一个更广泛的角度来看，网络本身可能就是一个合适的分析单位。例如，如果教师网络能够让思想和信息在课堂之间顺畅流动，那么就可以改善教育成果，因此一个连通性很好的管理者可以有效地协调课程改革。与此类似，一个二年级的老师知道很多关于刚上三年级的学生的信息，而这些信息对三年级的老师非常有帮助。数学老师知道学生还有什么概念没有掌握好，这类信息可以帮助科学老师更好地设计课程。因此，优秀的学校都拥有强大的教师网络。这只是一个例子；在许多其他公共领域和私人领域，网络都可以改进我们的思考。[17]

迈尔森值和结构洞

填补结构洞的人能够将网络中的不同社区联系起来，从而具有更大的影响力。网络的各种统计数据，例如介数，都与能不能占据结构洞相关。对于一个人在网络中的影响力的另一个衡量指标是迈尔森值，它依赖于夏普利值的原理。为了计算迈尔森值，我们在网络上构建一个合作博弈。在这个合作博弈中，只允许存在包含了连接组件的联盟。

考虑排成一行的三个人。假设他们的位置代表他们的政治（意识形态）立场。如图所示，B位于中间。如果我们限定只允许直接相邻的人组成联盟，那么位于最左的A将不能与位于最右边的C连接，除非B也加入联盟。为了计算出每个博弈参与者的迈尔森值，我们先为所有可行的联盟分配增加值；再对每个可能的联盟计算夏普利值，将每个联盟视为一个单独的博弈。最后，将每个联盟博弈的夏普利值相加，就可以得到迈尔森值。



可能的次序：**ABC**，**BAC**，**BCA**，**CBA**排除掉的次序：**ACB**，**CAB**

举例来说，假设两个博弈参与者的任何联盟产出的价值为10，

并且所有三个博弈参与者在一起时的产出的价值为14，那么就可以求出如下迈尔森值：博弈参与者A为3；博弈参与者B为8；博弈参与者C为3。[18]

中心性测度（例如介数）只以网络为基础，而迈尔森值则取决于价值函数。同时采用这两种测度，可以将一个人的权力对他在网络中位置的依赖性，与对他所发挥作用的依赖性分开。在这个例子中，三个博弈参与者的迈尔森值（3，8，3）与他们的介数得分（0，1，0）是完全相关的。但是情况并非总是如此，特别是对于更加复杂的网络和价值函数。

11　广播模型、扩散模型和传染模型

就像传染病的传染导致更多的传染病一样，信任的“传染”也可以促成更大的信任。

玛丽安·穆尔（Marianne Moore）

在本章中，我们运用广播模型、扩散模型和传染模型分析信息、技术、行为、信念和传染病在人群中的传播。这些模型在通信科学、市场营销学和流行病学的研究中发挥着核心作用。所有这三类模型都将整个人口划分为两个群组：知道或拥有某种东西的人与不知道或不拥有某种东西的人。随着时间的推移，个人会在这两个群组之间移动。有人会从易感者变为感染者，或者从不了解新产品新思想的人变成知情达意者。

串上某种传染病、购买了某种产品或掌握了某个信息的人的数量随时间推移而演变的曲线，也就是采用曲线的经验图形往往是凹的或S形的。人们如何获悉信息或怎样患上传染病——无论是通过广播传播还是扩散传播，决定了这种图形的形状。本章的主要内容就在于，将思想和传染病传播的微观过程与这些采用曲线的形状联系起来。为此，本章首先分析了广播模型。这种模型适用于人们从某个单一来源知悉思想或罹患传染病的情形。广播模型生成的图是r形的。然后，我们讨论扩散模型。在扩散模型中，传播始于接触，就像传染病在人与人之间传播时那样。扩散模型会产生S形曲线。

许多产品、应用程序、思想和信息都是通过广播传播和口碑传播的。我们可以同时允许广播和扩散来对这些情况建模，由此而得到的模型被称为巴斯模型（Bass　Model），它在营销学中起着核心作用。巴斯模型会生成r形曲线还是S形曲线，则取决于广播过程和扩散过程之间的相对优势。

本章中讨论的最后一个模型来自流行病学的关于传染病传播的SIR模型（易感者、感染者和痊愈者模型）。这个模型包括了痊愈率，它能够刻画抵御传染病的免疫系统、突显性逐渐弱化的信息，以及被流行浪潮甩下的时尚行为等。SIR模型会产生一个临界点，在临界点上，产品或传染病性质的微小变化，就意味着失败与成功之间巨大分野。病毒毒力的轻微减少，就可以使大规模传染变为轻微的发病；歌曲传唱概率的小幅变动，就可能把一只热门的新乐队送上天堂或打入地狱，区别之

大，恰如披头士乐队与一支在利物浦某个地下酒吧卖唱的乐队之间的区别。

广播模型

本章中介绍的所有模型都要假设存在一个相关人群，用N_{POP}表示。相关人群包括那些可能患上传染病、了解信息或采取行动的人。相关人群所指的并不是一个城市或国家的全部人口。如果我们要为连续主动脉缝合法的扩散建模，那相关人群就是指心脏外科医生，而不是居住在费城的所有人。

在任何时候，总会有些人患上了某种传染病、了解特定信息或采取了一定行动。我们将这些人称为感染者或知情者（用I_t表示），相关人群中除了感染者或知情者之外的其余成员则是易感者（用S_t表示）。这些易感者可能会感染传染病、了解信息或采取行动。[1]相关人群的总人数等于感染者或知情者人数加上易感者人数的总和：$N_{POP}=I_t+S_t$。

广播模型

$$I_{t+1}=I_t+P_{broad}\times S_t$$

其中，P_{broad}表示广播概率，I_t和S_t分别等于时间t上的感染者（知情者）和易感者的数字

初始状态为$I_0=0$，且$S_0=N_{POP}$。

广播模型刻画了思想、谣言、信息或技术通过电视、广播、互联网等媒体进行的传播。大多数时事新闻都是通过广播形式传播的。这个模型的目标是描述一个信息源传播信息的过程，可以是政府、企业或报纸。它也适用于通过供水系统传播污染的情况。但是，这个模型不适用于在人与人之间传播的传染病或思想。由于广播模型更适合描述思想和信息的传播（而不是传染病的传播），所以我们在这里说知情者的人数，而不说感染者的人数。

在给定时间段内，知情者人数等于前一期的知情者人数加上易感者听到信息的概率乘以易感者人数。按照惯例，初始人口全部由易感者组成。要计算未来某期的知情者人数，只需要将知情者人数和易感者人数代入上述方程即可。由此得到的将是一个r形采用曲线。

想象一下：某个拥有100万居民的城市的市长宣布了一项新的税收政策。在他宣布之前，没有人知道这项政策。假设某人在任何一天听到这个新闻的概率等于30%（即，$P_{broad}=0.3$），那么第一天会有30万人听到这个新闻。在第二天，剩下的70万人中有30%的人，即21万人会听到这个新闻。在每一个时期，知情者的人数都会增加，并且以一个递减的速度增加，如图11-1所示。

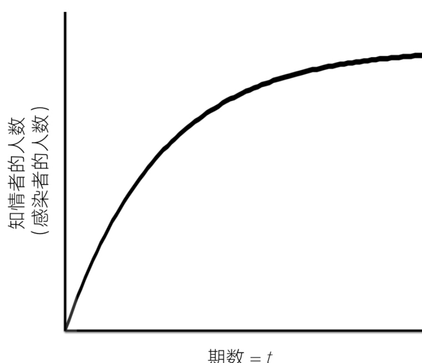


图11-1　广播模型产生的r形采用曲线

在广播模型中，相关人群中的每一个人最终都会知悉信息。如果有适当的数据，就可以估计出相关人群的规模。假设一家企业为练习太极拳的人推出了新设计的运动鞋，并在第一个星期就收到了20 000双鞋的订单。如果在第二个星期收到了16 000双鞋的订单，那么我们可以大致估计出他们最终的总销售量，也就是相关人群的规模为100 000。

用广播模型拟合销售数据

第1期：$I_1=20\ 000=P_{broad}\times N_{POP}$；

第2期：$I_2=36\ 000=20\ 000+P_{broad}\times（N_{POP}\text{-}20\ 000）$

于是，总销售量为[2]：$P_{broad}=0.2$，$P_{broad}=100\ 000$

当然，对于根据仅有的两个数据点估计出来的任何结果，我们都不应该抱以太大的信心。这个模型无疑遗漏了许多现实世界的特征。人们既可能通过传媒获悉相关消息，也可能通过口耳相传听到消息，而且有些人可能会购买不止一双鞋子，或者可能存在针对潜在消费者的广告，等等。如果把这些因素都包括进去，估计出来的结果肯定会有所不同。尽管必须牢记这个注意事项，但是这个模型确实提供了一个粗略的估

计。这个企业不应该期望能够卖出200万双鞋，但是应该有信心可以卖出不止100　000双鞋。随着更多数据的出现，估计结果是可以得到改进的。如果第三个星期的销售额是13　000双（这等于模型预测的数量），那么这个企业对当初的预测可以寄予更大的信心。

扩散模型

大多数传染病，以及关于产品、思想和技术突破的信息，都是通过口口相传而传播开来的，扩散模型刻画了这些过程。扩散模型假设，当一个人采用了某种技术或患上了某种传染病时，这个人有可能将之传递或传染给与他接触的人。在传染传染病的情况下，个人的选择不会在其中发挥任何作用。一个人患上某种传染病的概率取决于诸如遗传、病毒（细菌），甚至环境温度等因素。在炎热潮湿的季节，疟疾的传播速度要比在寒冷干燥的季节快得多。

技术的传播则与采用者的选择有关，因此更有用的技术被采用的概率更高。但是在这里，我们并没有在模型中明确将这种情况选择考虑在内。这样一来，苹果智能手表的新潮性就发挥了与流感病毒同样的作用。

在这里，我们更看重的是信息的传播，因此我们将人们分为知情者或不知情的新人。如果新人与知情者相遇且信息在他们之间传播，那么新人就会变成知情者。这种事件的发生，因环境而异。生活在城市中的人，相遇的概率可能比生活在农村的人更高，同时也有更高的接触概率。非常吸引人眼球的新闻也比一般的新闻被分享的概率更高，例如，关于外星人降临登陆的新闻被分享的概率比关于M&M公司的椒盐卷饼重新上市的新闻更容易被分享。因此，我们可以将扩散概率（diffusion probability）定义为接触概率（contact　probability）和分享概率（sharing probability）的乘积。我们可以根据扩散概率来构建模型，但是在估计或应用模型时，必须独立地跟踪接触概率和分享概率。

扩散模型假定随机混合（random　mixing）。随机混合的含义是，相关群体中任何两个人接触的可能性都相同。对于这个假设，我们应该保持警惕。就描述幼儿园内传染病传播的扩散模型而言，这可能是一个准确的假设，因为幼儿园里儿童之间的相互接触是高频率的。但是，如果将它应用于城市人口则是有问题的。在城市中，人们并不是随机混合的。人们在一定的社区中生活，在一定的场所内工作，他们属于工作团

队、家庭和社会团体，他们的互动主要发生在这些群体中。但是同时也不要忘记，一个假设要成为有用模型的一部分，其实不一定非得十分准确不可。因此，我们将继续使用这个假设，同时保持开放的心态，在需要改变的时候随时改变这个假设。

扩散模型

$$I_{t+1}=I_t+P_{diffuse}\times\frac{I_t}{N_{POP}}\times S_t$$

其中，$P_{diffuse}=P_{spread}\times P_{contact}$。

在这个模型中，与在传播模型中一样，从长期来看，相关人群中的每个人都会掌握信息。不同的是，扩散模型的采用曲线是S形的。最初，几乎没有人知情，I_0很小。因此，能够与知情者接触的易感者人数也必定很小。随着知情者人数的增加，知情者与不知情者之间接触的机会增加，这又使知情者的人数更快地增多。当相关人群中几乎每个人都成了知情者时，新知情的人数会减少，从而形成了S形的顶部。技术的采用曲线通常也具有这种形状。例如，杂交种子的采用曲线虽然因州而异（艾奥瓦州采用杂交种子的速度比亚拉巴马州更快），但是所有州的采用曲线都是S形的。[3]

在广播模型中，根据数据估算相关人群规模是一件相当简单的事情。采用者的初始数量与相关人群规模密切相关。与此相反，利用扩散模型的数据估计相关群体的规模可能会非常困难。产品销售量的增加，可能是由于一个很小的相关人群内部的高扩散概率，也可能是由于一个很大的相关人群中的低扩散概率。

图11-2显示了两个假想的智能手机应用程序的相关数据。在第一天，每个应用程序都有100人购买。在接下来的5天中，应用程序1拥有更高的总销量和更快的销量增长。如果没有模型，我们很可能会预测应用程序1拥有更大的市场。但是，用模型拟合这两组数据的结果表明，事实与我们猜想的恰恰相反。

上市日	智能手机应用程序 1	智能手机应用程序 2
1	100	100
2	136	130
3	183	169
4	242	220
5	316	286
	…	
365	1 000	1 000 000

图11-2 智能手机应用程序的两条采用曲线

应用程序1拟合的扩散概率为40%，相关人群规模为1 000人；而应用程序2的扩散概率为30%，相关人群规模为100万人。[4]事实上，只要再过几天，我们就会观察到应用程序2的相关人群更大。但是，如果没有模型，如果不能根据前5天的数据来进行分析，我们就可能会对总销售额给出不正确的推断。

在使用扩散模型来指导行动的时候，我们必须将扩散概率分解为分享概率和接触概率的乘积。为了提高应用程序的销售速度，开发人员既可以设法提高人们相互接触的概率，也可以设法加大他们分享关于应用程序信息的概率。要想改变第一个概率是很困难的。为了增大第二个概率，开发人员可以为带来了新注册用户的老用户提供一些激励，事实上，许多开发人员都是这样做的，比如游戏开发者可能会给带来了新注册玩家的老玩家奖励游戏积分。虽然这样做能够增加扩散速度，但是并不会影响总销量，至少根据这个模型来看不会有影响。如上所述，总销量等于相关人群的规模，而与分享概率高低无关，提高销售速度不会带来长期的影响。

大多数消费品和信息都是通过广播和扩散传播的。而巴斯模型则将这两个过程组合在一起了。[5]巴斯模型中的差分方程等于广播模型和扩散模型中的差分方程之和。在巴斯模型中，扩散概率越大，采用曲线的S形就越显著。电视、收音机、汽车、电子计算机、电话机和手机的采用曲线形状都是r形和S形的组合。

巴斯模型

$$I_{t+1}=P_{broad}\times S_t+P_{diffuse}\times\frac{I_t}{N_{POP}}\times S_t$$

其中，P_{broad}＝广播概率，$P_{diffuse}$＝扩散概率。

SIR模型

到目前为止，在我们已经讨论过的模型中，一旦有人采用了一项技术，则永远不会放弃它。对于电力、洗碗机和电视等技术来说，确实如此：一旦采用之后，一般永远不会不采用。但这并不适用于所有通过扩散传播的事物，例如我们患上了某种传染病之后不久就会恢复健康，或者当我们采用了某种流行款式或参加了某项潮流运动之后（例如，某种时装或舞蹈），是可以放弃的。遵循惯例，我们将放弃所采用的某种事物的人称为痊愈者。由此产生的模型，即SIR模型（易感者、感染者、痊愈者），在流行病学中占据了中心位置。

由于这个模型起源于流行病研究领域，同时也因为考虑传染病的痊愈更为自然，因此我们以传染病的传播为例来描述SIR模型。为了避免过于复杂的数学计算，我们假设治愈传染病的人会重新进入易感人群，也就是说治愈传染病并不会产生未来对传染病的免疫力。

SIR模型

$$I_{t+1}=I_t+P_{contact}\times P_{spread}\times\frac{I_t}{N_{POP}}\times S_t-P_{recover}I_t$$

其中，$P_{recover}$，P_{spread}，和$P_{recover}$分别等于传染病的传播概率、接触概率和痊愈概率。

流行病学家对接触概率和传播概率会进行单独跟踪，我们也会这样做。接触概率取决于传染病如何从一个人传播到另一个人。艾滋病通过性接触传播；白喉通过唾液传播；流感病毒通过空气传播。因此，流感的接触概率高于白喉，白喉的接触概率又高于艾滋病。而且，在发生接触后，各种传染病的传播概率也会有所不同。白喉比SARS更容易传染给另一个人。

SIR模型会产生一个临界点，就是所谓的基本再生数R_0，也就是接触概率乘以扩散概率与痊愈概率之比。某种传染病，如果R_0大于1，那么这种传染病就可以传遍整个人群，而R_0小于1的传染病则趋于消失。在这个模型中，信息（或者，在这个例子中是传染病）并不一定会传播

到整个相关人群。能不能做到这一点取决于R_0的值。因此，像疾病控制中心这样的政府机构必须依据对R_0的估计来指导政策制定。[6]

基本再生数R_0

$$R_0 = \frac{P_{spread} \times P_{contact}}{P_{recover}}$$

如表11-1所示，麻疹可以通过空气传播，因而它的再生数高于艾滋病，艾滋病只能通过性接触和共用针头传播。对R_0的估计假设人们不会为了应对传染病而改变行为。然而，当学校里虱子肆虐时，家长的反应可能是让孩子待在家中，以降低接触概率，还可能会剃光孩子的头发，减少接触发生时传播的可能性。这两种行为变化都会降低虱子传播的R_0。

表11-1　各传染病的基本再生数R_0

	麻疹	脊髓灰质炎	艾滋病	流感
R_0	15	6	4	3

在没有疫苗的情况下，检疫是一个选择，但是成本很高。[7]如果存在疫苗，那么疫苗接种可以预防传染病传播。即便做不到每个人都接种疫苗，也可以预防传染病传播。必须接种疫苗的人的比例，即疫苗接种阈值（vaccination threshold），可以通过公式$V \geq \frac{R_0 - 1}{R_0}$求出。我们可以从上述模型中推导出这个公式。[8]

疫苗接种阈值随R_0的增加而提高。例如，脊髓灰质炎的R_0为6，因此为了防止脊髓灰质炎的传播，疫苗必须覆盖5/6的人群。而麻疹的R_0为15，为了阻止麻疹的传播，疫苗必须覆盖14/15的人口。疫苗接种阈值的数学推导也为决策者提供了指引，如果接种疫苗的人数太少，这种传染病就会传播开来，因此政府接种疫苗的次数会超过模型估计的阈值。对于麻疹和脊髓灰质炎等R_0非常高的传染病，政府将努力保证所有人都接种疫苗。

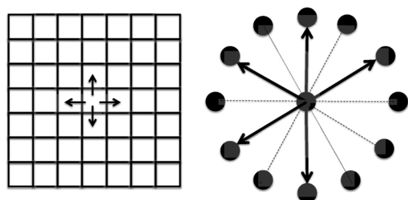
有些人担心疫苗有副作用，选择不参加疫苗接种计划。如果这些人只占人口的一小部分，那么其他人接种疫苗也可以防止这些人感染这种传染病，流行病学家将这种现象称为群体免疫力。选择不接种疫苗的人事实上是搭了其他接种疫苗的人的便车，对于搭便车的现象，我们将在

本书后面的章节中详细研究。[9]

R_0、超级传播者，以及度的平方

假设随机混合R_0的推导如下：在每一步，人群中的个体随机相遇。如前所述，随机混合假设可能与通过空气传播的传染病或通过接触传播的传染病有关，但对于通过性行为传播的传染病则不太合理。

如果将SIR模型嵌入到网络中，就会观察到度分布对传染病传播的重要性。在这里，我们比较一下矩形网格网络（棋盘格）与中心辐射型网络。在矩形网格网络中，每个节点都连接到东、南、西、北的节点；而在中心辐射型网络中，则由一个中心节点连接到所有其他节点。



假设传染病会随机发生在某个节点上。我们在网络中设定$p_{contact}=1$，以保证每个人都会与他所连接的每个人接触。在下一个时期，传染病可能以一个与传染病毒力相对应的给定概率，独立地扩散给每个邻居。

首先考虑矩形网格网络。在每个时期，传染病都可以扩散到东、南、西、北4个节点中的任何一个。我们预计，如果传染病传播的概率超过了1/4，传染病就会蔓延。展望未来的一个时期就会看到，如果一个新节点患上了传染病，那么这个节点有3个可能患上传染病的邻居。如果原始节点的北部和东部的两个邻居患上了传染病，那么传染病可能传播的节点就会达到6个。因此，这种网络似乎对传染病传播的潜力发挥没有太大影响。

接下来，考虑中心辐射型网络。第一个患上传染病的节点可能是中心节点，也可能是外围节点。如果中心节点患上了某种传染病，那么它可以将传染病传播到任何一个其他节点。我们预计这种传染病会扩散，即便传播的概率很低也是如此。如果是一个外围节点患上了传染病，那么唯一可能被传染的节点就是中心节点。正如

在前面讨论过的，如果中心节点患上了传染病，那么即使传播的可能性很小，传染病也会蔓延。

对于中心辐射型网络，R_0携带的信息量很有限，因为如果中心节点患上了传染病，传染病就会传播开来。流行病学家们将位置在度很高的中心节点上的人称为“超级传播者”（superspreaders）。超级传播者加速了艾滋病和SARS的早期传播。[10]超级传播者不一定是社交明星或“人脉”特别广的人，可能从事某种特定的行业职业，比如收费站的收费员、银行柜员、牙科医生，这类职业使他会与属于不同社交网络的人接触。生活在19世纪与20世纪之交的“伤寒玛丽”（Typhoid Mary）只是纽约的一名上门服务的厨师。她从这一家再到另一家，将伤寒感染给每一个接触者。当她被确认为传染源之后，就被强制隔离了。

为了推导出高度数节点的影响，我们首先要注意到一个事实：高度数节点不但能够更快地传播传染病，而且会更快地患上传染病。如果一个人朋友的数量是另一人的三倍，那么他患上传染病的可能性也是后者的三倍，同时传播这种传染病的可能性也是后者的三倍。因此，他对传染病传播的总贡献将是另一个人的九倍。因此，节点对传染病（或思想）传播的贡献与节点的度的平方相关。如果节点A的度数是节点B的K倍，那么节点A传播传染病的可能性是节点B的K倍，同时传染病传播到节点A的概率也是节点B的K倍。因此，节点A对传染病传播的总效应将是节点B的K^2倍。这种现象被称为度的平方。

一对多

尽管SIR模型原本是用来分析传染病传播的，但是我们也可以将它应用于所有先通过扩散传播，然后趋于消失的社会现象，例如书的销售、歌曲的流行、舞步的风行，“热词”的传播、食谱和健身方法的流传等。在这些情形下，我们也可以估计接触概率、传播概率和“痊愈”概率，以及基本再生数R_0。这个模型意味着，这些概率只要发生了微小的变化，就可以使R_0移动到高于零的水平，从而造成成功与失败之间的天壤之别。

确实，成功可能取决于非常微小的差异，正如美国作家约翰·厄普代克（John Updike）在描述棒球明星特德·威廉姆斯（Ted Williams）最

后一次击球时所说的："一件事情做得很好与搞砸了之间，只有极其细微的差异。"[11]假设你构思了一个新的笑话；只要让这个笑话更有趣一点点，就可能会把R_0推到高于1的水平，从而使这个笑话广泛传播开来。同样的逻辑也适用于想法的"黏性"。如果一个想法能够在人们的思维中再坚持一小段时间，那么他们摆脱它的"痊愈"率就会降低，从而提高了R_0。

当然，并不是所有情况都会位于阈值上。披头士乐队拥有巨大的才华，他们的R_0肯定超过了1，尽管这只是一个猜想。对于现在的流行歌星，我们可以使用互联网下载量来估计他们的R_0。流行歌星贾斯汀·比伯（Justin Bieber）的R_0估计为24，这就是说，他的传染"毒力"比麻疹更强。[12]

在SIR模型中，我们推导出了两个关键阈值，即R_0和疫苗接种阈值。这两个阈值都是属于敏感依赖于环境的临界点，环境（情境）中的微小变化都会对结果产生很大的影响。这种临界点不同于直接临界点（direct tipping point）。在直接临界点，特定时刻的微小行动会永久性地改变系统的路径。直接临界点出现不稳定的点是，例如当球停在山顶上时。在任一方向上稍微推一下，都将会使球从山的这一侧或另一侧滚下去，这个小小推动是一个直接的倾覆。[13]

而在依赖于环境的临界点上，参数的变化会改变系统的行为方式。在直接临界点上，未来的结果轨迹急转直下。折弯，例如由扩散模型产生的*S*形采用曲线中的第一个弯曲，不满足这两种临界点的定义。采用曲线中的折弯对应于斜率增长率最大的点。在那一点上，扩散一发不可收拾，但是并没有发生倾覆。

图11-3显示了Google+发布后前两个星期的用户数。[14]从图中可以看出，在发布6天后，出现了一个折弯。在那一点上，扩散的过程正在顺利展开。从两个星期内就获得了超过1 600万用户这个结果来看，我们不能说Google+很早就陷入了困境，更不能说它在第6天就出现了直接临界点。将倾覆与急剧上升（下降）混淆起来，导致临界点这个术语被过度滥用了。新闻媒体和互联论坛上所说那些临界点，几乎有很少符合正式定义的。

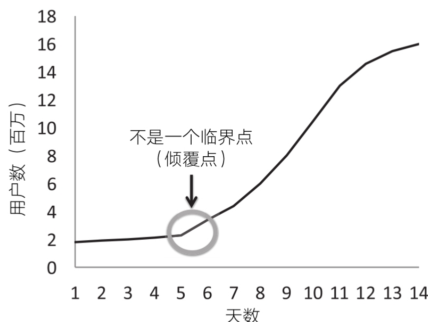


图11-3　Google＋用户数量上的一个折弯点（不是一个倾覆点）

我们不妨将肥胖症视为一种流行病来考虑。尽管人们不会像患上感冒那样感染肥胖症，但是他们可能会受到某种社会影响而做出一些容易导致肥胖的行为。[15]要想扭转肥胖这种流行病，我们必须降低它的R_0。而要降低R_0，则可以通过降低接触概率或者提高分享概率和痊愈概率来实现。当然，在很多方面，用SIR模型来研究肥胖症的传播、学校辍学率或犯罪率，并不比经济学模型或社会学模型更好。它只是一个不同的模型，因此会给出不同的解释和预测，它也可能指向不同的行动或政策。它扩大了我们的模型集合，帮助我们更好地理解世界，但它不是解决问题的灵丹妙药。

在将广播模型、扩散模型和传染模型应用于社会现象时，我们可能会发现某些假设是成立的，而其他一些假设则不能成立。例如，在某种传染病的传播中，每一次接触导致该传染病传播的概率是独立的。但是在社交领域，由于采用本身也是一种选择，因此传染有可能会因更多的接触（曝光）而变得更有可能。流感不是我们选择的，我们只是得了流感。但是我们会选择买紧身牛仔裤，随着越来越多的人穿上了紧身牛仔裤，我们所有人都更可能穿紧身牛仔裤。类似的逻辑也适用于分析社交运动的参与率、新技术的采用率，甚至分析文身的人有多少。在这些情况下，我们可能必须对基本模型进行修正，以允许每次接触的采用概率会随着接触次数的增多而增大。[16]信仰或信任行为的“传染”也是如此。这种修正，在扩大模型的应用范围时通常是必不可少的。

12　熵：对不确定性建模

信息是不确定性的解。

克劳德·香农（Claude Shannon）

在本章中，我们讨论熵。熵是对不确定性的一个正式测度。利用熵，我们可以证明不确定性、信息内容与惊喜之间的等价性。低熵对应于低不确定性，同时揭示的信息很少。如果某个结果发生在低熵系统中，例如太阳从东方升起，我们并不会感到惊讶。而在高熵系统中，比如在抽奖时抽中了某个数字，结果是不确定的，并且实现的结果能够揭示信息。在这个过程中，我们经历了惊喜。

利用熵，可以比较不同的现象。我们可以判断新西兰的选举结果是不是比联合国对谴责某个国家的方案的投票结果更不确定，还可以将股票价格的不确定性与体育赛事结果的不确定性进行比较，也可以利用熵的概念来区分四类结果：均衡、周期性、复杂性和随机性。我们可以将看似随机的复杂模式和真正的随机性区分开来，并且可以分辨出哪些现象看起来像是有一定模式的，但事实上是随机的。

我们还可以使用熵来表征分布。在没有控制或调节力量的情况下，一些群体可能会向最大熵漂移。给定特定的约束条件，例如不变的均值或方差，就可以解出最大熵分布。最大熵分布的结果还可以用来证明某些分布比其他分布更优，从而能够对我们在建模时的选择起到指导作用。

本章分为五个部分。在第一部分中，我们讨论对信息熵（information　entropy）的直观认识，然后给出信息熵的正式定义。在第二部分中，我们描述了关于熵的公理基础。在第三部分中，我们讨论如何使用熵来区分均衡、周期性、随机性和复杂性。在第四部分中，我们研究了会在给定约束条件下产生最大熵的系统。最后，我们探讨了这样一个问题：为什么在有的时候，我们更喜欢复杂性而不是均衡。

信息熵

熵是用来度量与结果的概率分布相关的不确定性的。因此，它也可

以衡量意外。熵与方差不同，方差度量一个数值集合或数值分布的离散程度。不确定性与离散程度有关，但是两者并不是一回事。在具有高不确定性的分布中，许多结果的概率都是有意义的，这些结果并不一定有数值，具有高离散度的分布则只是具有一些极端的数值。

通过比较具有最大熵的分布与具有最大方差的分布，可以将这种区别鲜明地呈现出来。给定取值范围为从1到8的整数的若干结果，能够使最大化熵的分布对每个结果赋予相同的权重。[1]而能够使方案最大化的分布则是以1/2的概率取值1、以1/2的概率取值8（图12-1）。

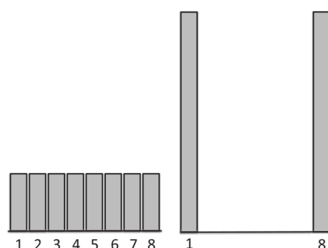


图12-1　最大熵与最大方差

熵是在概率分布上定义的。因此它可以应用于非数值数据分布，例如森林中鸟儿的种类或不同口味果酱的市场份额。熵在数学上等于概率与它们的对数之和的相反数。这个数学公式听起来似乎很复杂，但是事实并非如此。

我们先从信息熵这种特殊情况开始讨论。对于信息熵，可以把它理解为根据随机抛硬币的结果来衡量不确定性的一种方法。假设每个家庭都只有两个孩子，男孩和女孩的可能性相同。某个家庭的孩子们的性别列表（按出生顺序排列）相当于抛两次硬币。因此，结果分布的信息熵为2，因为它对应于两个随机事件。其信息内容也等于2，因为我们只要提出两个“是或否”问题要求他们回答，就可以掌握结果。

与此类似，在有三个孩子的家庭中，性别列表相当于抛3次硬币。要了解这样的家庭的孩子的性别，也只需要提出三个“是或否”问题。同样的逻辑适用于任何数量的儿童。在一般情况下，要了解N个孩子的性别，只需要提出N个“是或否”问题。

这里需要注意的是，这N个问题区分出了2^N种可能的出生顺序。这种数学关系是理解熵测度的关键所在：N个二元随机事件会产生2^N个可能的结果序列，并且，与之等价，我们可以通过提出N个“是或否”问题知悉结果序列。这也意味着，信息熵将不确定性水平（和信息内容）N

分配给了2^N个结果上的一个等可能分布。

于是，挑战变成了如何用数学公式刻画这种关系。每个结果序列的概率均为$\frac{1}{2^N}$。要将这个数值转换为N，需要一个相当复杂的数学公式$N=-\log_2(\frac{1}{2^N})$。[2]我们可以将这个公式推广到任意概率的情形下。如果某个结果序列出现的概率为p，就分配一个不确定性$\log_2(p)$，它近似于识别该序列所需要提出的“是或否”问题的数量。为了计算出一个分布的信息熵，我们只需求得所有结果（或者像在前面那个例子中那样的结果序列）需要提出的问题的期望数量的平均值。

信息熵

给定一个概率分布（p_1，p_2，…p_N），信息熵，H_2等于：

$$H_2(p_1,p_2,\ldots p_N)=-\sum_{i=1}^{N}p_i\log_2(p_i)$$

注：上面的下标2表示使用的是以2为底的对数。

乍一看，这个数学公式带来的混淆似乎比它所能澄清的还要多。通过举例说明，应该能够使这个公式更加直观。想象一下这种情况：第一胎是女孩的家庭不再生任何孩子，而第一胎是男孩的家庭则还要再生两个孩子。从而，所有家庭中将有一半家庭只有一个女孩。而另一半家庭则等可能地分别属于如下四个结果之一：三个男孩；两个男孩与一个女孩；一个男孩与两个女孩；一个男孩一个女孩与一个男孩。这四种结果中的每一种出现的概率均为1/8。

信息熵等于我们想了解一个家庭的子女排列状况时必须提出的“是或否”问题的期望数量。我们首先会问，第一个孩子是不是女孩，回答“是”的概率为1/2。如果是这个答案，那么就不需要继续问下去了。因此有一半的时间，我们只需问一个问题。我们可以把这写成$-\frac{1}{2}\log_2(\frac{1}{2})$。如果答案是否定的，那么我们还必须再提出两个问题，于是总共要问三个问题。这四种情况中的每一种都以1/8的概率出现，因此每种情况对信息熵的贡献为$1/8\times3$，我们对每种情况可以写出$-\frac{1}{8}\log_2(\frac{1}{8})$。从而，信息熵等于2，即上述五项的总和。[3]虽然这里使

用的对数和负号可能会让有些人觉得困扰，但是直观含义仍然是非常清楚的：信息熵就对应着“是或否”问题的期望数量。如果我们不得不提出很多问题，那么分布就是不确定的。而知道了结果，也就揭示了信息。

熵的公理基础

公理基础：熵

$H_a(p_1,p_2,\ldots p_N) = -\sum_{i=1}^{N} p_i \log_a(p_i)$，其中，$a > 0$。

这种熵测度是唯一满足以下四个公理的测度：

对称性：对于任何转置概率，都有连续函数$H(\sigma(\vec{p})) = H(\vec{p})$。

最大化：对于所有N，$H(\vec{p})$在$p_i = \frac{1}{N}$处最大化。

零性：H（1，0，0，…0）＝0。

可分解性：如果$\vec{P} = (p_{11}, p_{12} \cdots p_{nm})$

$$H(\vec{P}) = H(P_1, P_2, \ldots P_N) + \sum_{i=1}^{N} H(Q_{P_i})$$

其中，$P_i = \sum_{j=1}^{m} p_{ij}$ 且 $Q_{P_i} = \left(\frac{p_{i1}}{P_i}, \frac{p_{i2}}{P_i}, \ldots, \frac{p_{im}}{P_i}\right)$。

为了得到熵的一般表达式，我们采用公理化的方法。数学家克劳德·香农对他给出的这种测度施加了四个条件。前三个条件很容易理解，它必定是连续的和对称的，而且在所有结果以相同的概率发生时最大化，同时在某些结果上等于零。第四个条件可分解性则要求在具有m个子类别的n个类别上定义的概率分布的熵，等于各类别上的分布的熵与每个子类别的熵的总和。香农证明，有一类熵测度是唯一满足这些公理的测度。虽然这里的分布的乘积是一个不那么直观的自然假设。例如，在结果是两个独立事件的乘积的情况下，这意味着联合事件的信息内容是每个事件单独发生时的信息内容的总和。

正如夏普利值的公理基础一样，这些公理对存在性的贡献大于它们本身的合理性。聪明的数学家总是可以构造出能唯一定义一个函数的公理。香农的前两个公理很难质疑。有的人可能吹毛求疵地指责，将已知分布的不确定性设置为零过于任意了，但这只是一个适当的基准，另一

种可能性是将已知分布的不确定性指定为1。[4]可分解性虽然解释起来不是很容易，但是也很难去挑战它。两个组合随机事件的不确定性理应等于每个事件的不确定性之和。总的来说，这些公理不仅仅是可辩护的，事实上，它们是难以辩驳的。

利用熵区分结果类别

我们现在阐明，如何利用熵测度来对经验数据进行分类，并在计算机科学家、数学家斯蒂芬·沃尔弗拉姆（Stephen Wolfram）给出的四大类别的框架下建模：均衡、周期性、随机性和复杂性。[5]在沃尔弗拉姆的这个分类中，放在桌子上的铅笔处于均衡状态，绕太阳运转的行星处于循环当中，抛硬币的结果序列是随机的，纽约证券交易所的股票价格也是近似随机的（我们在下一章中将会说明原因）。最后，一个人大脑中的神经元发放则是复杂的：它们既不会随意发放，也不会以某个固定的模式发放。图12-2以图形方式呈现了这四个类别。

平衡结果没有不确定性，因此其熵等于零。周期性过程具有不随时间变化的低熵。当然，完全随机过程具有最大的熵。复杂性具有中等程度的熵，因为复杂性位于有序性和随机性之间。虽然熵在两种极端情况下能够为我们给出明确的答案——均衡和随机性；但是这并不适用于周期性和复杂性的结果。在这些情况下，通常还必须善用我们的判断力。

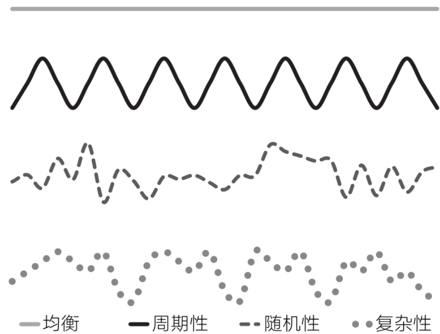


图12-2　沃尔弗拉姆的四种类别

为了对时间序列数据进行分类，我们需要先计算出不同长度的子序列中的信息熵。假设，有个人会把他每天戴的帽子的类型一一记录下来。假设他只在两种帽子之间进行选择，一种是贝雷帽，记为B，另一种是浅顶软呢帽，记为F。这样过了一年，他对帽子的选择生成了一个有365个事件的时间序列。我们先计算长度为1的子序列的熵，也就是

说，先计算戴每种类型帽子的概率的熵。假设他喜欢这两种类型的帽子的程度相同，那么长度为1的子序列的熵等于1。因此，我们可以先把均衡排除掉，因为他会改变他的选择，但是其他三种类别中的任何一种都是可能的。

为了确定类别，我们接下来计算长度为2到6的子序列的熵。如果所有都具有最大的熵，那么我们可以将简单的周期性排除掉。假设当我们考虑更长的序列时，熵会缓慢增加，直到达到最大值8为止。换句话说，无论子序列有多长，熵都不会超过8。熵为8相当于256个结果的等可能分布，这不可能是一个简单的循环。熵为8更可能代表具有特定结构和模式的复杂过程序列。我们不能确定地说，这个时间序列是复杂的。一种可能的情况是，这个人试图做到随机化，但是却失败了。

最大熵和分布假设

在很多情况下，我们建模时都必须把不确定性包括进来；因而作为建模者，必须对有关的分布做出假设。这里的原则是，我们要尽量避免做出任意特殊假设（ad hoc assumption）。也许，我们对产生分布的过程已经有了一些了解。如果是这样，通常可以运用逻辑—结构—功能方法，推导出该过程产生的统计结构。

例如，假设我们想要对一个房地产拍卖中的所有拍卖对象的总价值的分布做出一个假设。总价值等于各个项目的价值总和。因此，我们可以根据中心极限定理假设这是一个正态分布。对于一栋房子的可能价值，我们也可以假设一个正态分布，因为房屋的价值取决于它的多个性质：卧室的数量、浴室的数量和占地大小等。

对于艺术珍品或稀有手稿的可能价值，正态分布却可能没有意义。在这些情况下，我们对决定它们价值的过程几乎一无所知。一种方法是假设一个具有最大不确定性的分布，即最大熵分布。

最大熵分布的形状取决于各种约束条件。正如我们已经看到的，如果假设了一个最小值和一个最大值，那么均匀分布会使熵最大化。教科书和学术期刊中的许多社会科学模型都假设均匀分布，我们可能会质疑这个假设，因为均匀分布在现实世界中确实很少出现。然而，无差别原则（principle of indifference）可以证明假设均匀分布的合理性。如果只知道范围或可能集，那么就应当予以无差别的对待。

在某些情况下，我们可能知道分布的均值，也知道所有值都必定是正数。给定这些约束条件，最大熵分布必定具有长尾，因为我们要将分布置于更多的值上，从而必须使少数高值结果与许多低值结果保持平衡。不难证明，熵最大化分布是一个指数分布。因此，如果我们正在构建一个模型，需要假设网站点击量或市场份额的分布形式，那么在没有可用数据的情况下，指数分布是一种自然的假设。

如果我们确定了均值和方差（并且允许出现负值），那么最大熵分布则是正态分布。这里的逻辑与前一种情况类似。为了创造更多的不确定性，我们创造了一些极端值，在这里，可以平衡正值和负值，而不用改变均值。但是，这样做会增大方差，因此我们必须在均值附近添加更多值，从而创造出钟形曲线。

我们可以在逻辑—结构—功能框架内解释这些最大熵分布。如果我们认为在给定的社会、生物或物理环境中，某个微观层面的过程能够最大化熵，那么我们应该期待上面这些分布中的某一个会出现。或者也可以假设一个微观过程，并能够证明熵在增加。如果是这样，上述分布中的某一个也会涌现出来。

最大熵分布

均匀分布：给定范围［a，b］，使熵最大化。

指数分布：给定均值μ，使熵最大化。

正态分布：给定均值μ和方差σ^2，使熵最大化。

我们也可以将这些结果解释为探索性的。我们可能会遇到一些指数分布或正态分布的数据。虽然没有“义务”去追问某种潜在的行为是否会在一定约束条件下使熵增加，但这样做确实可以帮助我们获得一些新的洞见。在本书前面的章节中，我们利用中心极限定理解释了物种的高度、重量和长度为什么会服从正态分布。在这里，我们再给出一个不同的、基于模型的解释：如果一种突变能够最大化熵（以便探索最好的生态位），并且假设平均规模和总离散度是固定的，那么规模的分布将会是正态的。关键不在于这种最大熵方法是不是提供了一个更好的解释，而在于给定约束下最大化熵必定会导致正态分布。因此，当我们看到正态分布时，它可能是最大化熵的结果。

熵的实证含义和规范含义

前面我们已经讨论了，熵如何衡量不确定性、信息和惊喜，如何与测量离散度的方差不同，以及如何有助于我们对不同类别的结果进行分类和比较。在本书第13章和第14章中研究随机游走和路径依赖时，还会利用熵来识别随机性并测量路径依赖的程度。事实上，我们可以将熵测度用于任何实际应用，可以用它来衡量对金融市场的干预是增加了还是减少了不确定性，可以检验选举、体育赛事或博彩中的结果到底是不是随机的。

在这些应用中，熵都是作为一个实证的衡量标准来使用的。它告诉我们世界是什么样的，而不是世界应该是什么样。一个系统中的熵的本质，不能简单地说好，也不能简单地说不好。我们想要多少熵，取决于具体情况。在制定税法时，我们可能需要一种均衡行为模型，并不希望有随机性。在规划城市时，我们可能会希望看到复杂性，均衡或者周期性都会显得过于平淡。我们希望一个城市充满生机活力，为偶然的相遇和互动提供无限机会。在这种情况下，更多的熵会更好，但是又不能太多。我们不喜欢随机性，随机性会使计划变得非常困难，并可能导致我们的认知能力崩溃。最理想的情况是，世界会产生适度的复杂性，以保证我们生活在一个有趣的时代。

建筑师克里斯托弗·亚历山大（Christopher Alexander）证明，诸如强中心、厚边界和非独立这类的几何属性，能够生成复杂的生活建筑、社区和城市。[6]亚历山大渴望城市和生活空间中的复杂性。中央银行的规划者可能不太喜欢复杂性，在金融市场中，他们可能更喜欢可预测的均衡结果。不过幸运的是，使用模型，我们既可以探索复杂性，也可以讨论均衡的可能性。

13　随机游走

> 醉鬼能找到回家的路，但是一只醉酒的小鸟可能永远回不了家。
>
> 角谷静夫（Shizuo Kakutani）

在本章中，我们讨论两个来自概率论和统计学的经典模型：伯努利瓮模型（Bernoulli urn model）和随机游走模型。[1]这两个模型都描述了随机过程，即使看上去它们似乎在生成某种复杂的结构。如果不收集数据，随机性是很难辨别的。我们经常想当然地以为能够在选举结果、股票价格和体育赛事得分中总结出一定的模式，但这只是一厢情愿。借用学者、风险分析师纳西姆·塔勒布（Nassim Taleb）的一句俏皮话来说，我们都被随机性所惑，不过是一些“随机漫步的傻瓜”！[2]

伯努利瓮模型描述了产生离散结果的随机过程，例如抛硬币或掷骰子。这个模型在几个世纪以前出现时，是为了解释赢得赌注的概率，现在已经在概率论中占据中心位置。随机游走模型就是建立在伯努利瓮模型的基础上的，保持了正面和反面的总数。这个模型可以刻画液体和气体中粒子的运动，动物在物理空间中的活动，以及从出生到童年人体身高的增长，等等。[3]

本章首先简要介绍伯努利瓮模型，并对条纹长度（length of streaks）进行了分析。然后描述随机游走模型，我们将会了解到，一维和二维随机游走会无限次地回到起点，而三维随机游走则可能完全不需要回到起点。我们还会了解到，对于一维随机游走，回到零点之间的时间间隔分布遵循幂律分布。对于这个发现，有人可能认为它除了满足人们的好奇心之外没有什么用，但事实上，它可以解释物种和企业的生命周期。我们还将使用随机游走模型评估有效市场假设，并用它来确定网络规模。

伯努利瓮模型

伯努利瓮模型由一个装了灰球和白球的瓮组成。从瓮中抽取的球代表随机事件的结果。每次抽取都与之前和之后的抽取无关，因此我们可以应用大数定律：从长远来看，抽出每种颜色的球的比例将会收敛到这

个球在瓮中的比例。当然，这并不意味着从一个装了7个白球和3个灰球的瓮中抽取1 000次，将会恰好抽出700个白球，它的意思是抽取出来的白球比例会收敛到70%。[4]

伯努利瓮模型

每一次，从一个装了G个灰球和W个白球的瓮中随机抽取一个球，结果等于抽取出来的球的颜色。在下一次抽取之前，球要先放回瓮中。令$P=\frac{G}{(G+W)}$表示灰球的比例。在抽取N次的情况下，可以计算出抽取出来的灰球的期望数量N_G，及其标准差σ_{N_G}：

$$N_G=N\times P，以及，\sigma_{N_G}=\sqrt{N\times P\times(1-P)}$$

伯努利瓮模型的结果产生了可预测长度的条纹。在灰球和白球数量相等的瓮中，抽取出白球的概率等于1/2，连续抽取出两个白球的概率等于1/2乘以1/2，以此类推。一般情况下，如果瓮中白球的比例为P，那么连续抽取N个白球的概率等于P^N。通过计算概率，我们可以评估某种条纹是不是有可能出现（尽管很令人吃惊），或是几乎完全不可能（因而基本上可以肯定“有诈”）。当一名篮球运动员连续9次投中了三分球时，只是有热手效应吗？或者，我们是否应该期待有一个这种长度的随机序列？数学计算表明，一个很优秀的三分投手在长达10年的职业生涯中，也几乎完全不可能连续9次投中三分球。[5]

我们可以进行类似的计算以确定投资者是幸运、能力出众还是在欺诈。自1965年至2014年，由沃伦·巴菲特（Warren Buffett）经营的集团伯克希尔哈撒韦公司（Berkshire Hathaway），在50年中有42年的表现优于市场。1964年伯克希尔哈撒韦公司的1美元在2016年的价值已经超过了1万美元，而投资标准普尔500指数的1美元价值大约为23美元。如果伯克希尔哈撒韦公司有50%的机会击败市场，那么它在50年来的表现应该超过市场的25倍，标准差为3.5年$\left(3.5\approx\sqrt{50\times\frac{1}{2}\times\frac{1}{2}}\right)$。而事实上，伯克希尔哈撒韦公司击败市场的实际年数大约高于均值四个标准偏差，这是一个概率仅有百万分之一的事件，因此，我们可以排除这完全是运气的可能。由于伯克希尔哈撒韦公司定期公布它的投资，所以也可以排除欺诈的可能。与此相反，前纳斯达克主席、美国历史上最大的诈骗案制造者伯纳德·麦道夫（Bernard Madoff）从来不透露他的投资情况，如果客户要求投资透明度的话，麦道夫是不可能连续几十年取得“成功”，连

续几十年得到正回报的。[6]

随机游走模型

接下来讨论简单随机游走模型，它建立在伯努利瓮模型的基础上，并将过去结果的和保持下来。我们将初始值，也就是模型的初始状态设置为零。如果我们抽取出一个白球，就在总数上加1；如果抽取出一个灰球，就从总数中减1。模型在任何时候的状态都等于先前结果的总和，也就是抽取出来的白球总数减去抽取出来的灰球总数的值。

简单随机游走模型

$$V_{t+1}=V_t+R(-1,1)$$

其中，V_t表示时间t上的随机游走值，$V_0=0$，$R(-1,1)$是一个可能等于-1或1的随机变量。在任何时间段内，这个随机游走的期望值都等于零，且标准差为$\sqrt{t}$，其中的t等于周期数。[7]

图13-1给出了一个简单随机游走。这幅图看上去似乎有一个模式：先是一个长期下降的趋势，然后是一个上升趋势；在上升过程越过零线时出现了一个适度的崩溃。但这个模式只是偶然发生的。

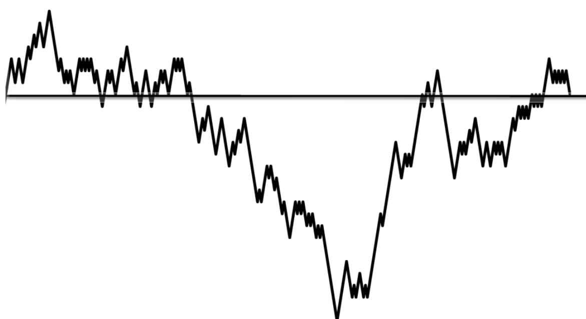


图13-1　一个300周期的简单随机游走

简单随机游走既是周期性的（会无限次地返回零点），又是无界性的（会超过任何正的或负的阈值）。如果等待足够长的时间，随机游走会高于正的1万、低于负的100万，也会无限次地穿过零线。此外，返回零点所需的步数分布满足幂律。[8]在大多数时候，返回零只需几步。所有游走中，有一半是两步返回的，然而有些游走需要很长时间才能返回。鉴于随机游走的无界性，这必定是真的。一个超过100万步阈值的游

走，需要超过200万步才能到达那里并返回零点。

幂律分布结果还有一个意想不到的应用领域。如果我们将企业的销售水平或员工规模建模为随机游走，那么企业的生命周期就会成为一个幂律分布。更准确地说，当销售强劲时，企业会新招聘一名员工；当销售不佳时，会解雇一名员工；当不再拥有任何员工时，企业也就“寿终正寝”了。这样一来，返回次数的分布就等于企业生命周期的分布，而且是一个幂律分布。再者，就其第一近似而言，企业的生命周期是一个幂律。[9]我们可以应用相同的逻辑来预测生物分类单元（界，门，经，纲，目，科，属和种）的寿命。如果某个分类单元的成员数量遵循随机游走，例如，如果某个属中的物种数量随机地上下变化，那么，这个分类单元的大小就应该满足幂律。这方面的数据支持了这个模型的预测。[10]

对于随机游走模型，还可以做这样一个类比：将随机游走视为冰川沿着地面的移动。根据模型的预测，冰川湖泊的大小分布将满足幂律。每一次，当冰川落到了陆地表面以下又返回顶部时，就会形成一个直径等于返回时间的湖泊。在这里，相关数据再一次与模型基本对应。[11]

这个基本随机游走模型可以通过多种方式加以修正。我们可以创建一个正态随机游走（normal random walk）。在正态随机游走中，每一周期的值的变化都服从正态分布。正态随机游走不会完全回到零点，但它会无限次地穿过零点。

我们还可以令某一种结果比另一种结果更有可能发生，从而创建一个有偏差的随机游走。我们可以利用这种有偏差的随机游走模型来预测在博彩中获胜的概率。轮盘赌中，在红色结果上下注时赢的概率等于9/12。[12]我们可以将赌轮盘赌的总收益或总损失建模为这样一个随机游走：增加1的概率为9/19（大约47.4%），而减少1概率则为10/19。那么在下注100次之后，预期损失为5美元，标准差为10美元。这也就是说，我们可以在95%的置信水平上，认为损失不超过25美元、收益不超过15美元。在下注1万次之后，预期损失等于526美元，标准差为100美元。因此，在95%的置信水平上，我们的损失介于325美元与725美元之间。[13]同样，在下注1万次之后，我们还能赢是一个相当于超过均值5个标准偏差的事件，也就是说我们赢的可能性不到百万分之一。因此，要想在轮盘赌中赢，应该做的事情是下一个大赌注而不是下很多个小赌注。

一些体育比赛，例如篮球比赛，可以建模为两个有偏差的随机游走。在球场上，每支球队在每次攻守中都有可能得分。这个概率可以根

据一支球队的进攻能力和对方球队的防守能力来估计。我们将球队在球场上的“行程”模拟为一个随机事件。每支球队的得分对应一个随机游走值，得分较高的球队更有可能获胜。来自NBA的数据分析表明，实际比赛结果与这个模型匹配得相当好。只有当一支球队获得了巨大的领先优势时，得分才会偏离随机性，在那种情况下领先优势继续扩大的可能性低于领先优势缩小的可能性。这种现象可以解释为领先的球队失去了继续得分的动力，同时落后的球队则必须至少让分数看上去不那么“丢脸”。[14]

我们似乎会认为篮球比赛的结果肯定不是随机的。聪明、健壮且灵活的篮球运动员，拥有很多巧妙的进攻手段，并能在关键时刻实现扭转乾坤的得分。这当然也是事实，但是球员们的努力效果可能会被抵消。额外的进攻得分可能会因为额外的防守努力而被抵消。一个重要的抢断后的快速上篮，可能会被冲刺了大半个球场的对方球员破坏。这个模型还提出了一个策略：更强的那支球队应该加快比赛节奏，以创造更多的进攻回合。占有优势的球队应该更频繁地玩“轮盘赌”，因为随机“漂移”对他们有利。

简单随机游走模型只在一个维度上进行。我们还可以对高维随机游走建模。二维随机游走从平面中的原点（0，0）开始，然后在每个周期中随机走向东、南、西、北。二维随机游走类似于在一张纸上绘制出来的一条弯弯曲曲的线，同时也满足递归性（recurrence）和无界性，有点儿类似于在你的起居室中随机搜索一只丢失的耳环时的路线。这种递归性使随机觅食成了蚂蚁的一个觅食策略。[15]如果二维随机游走不是递归性的，那么蚂蚁就需要更复杂的内部地图或更强的信息踪迹才能找到它们的巢穴。

但是在有三个维度的情况下，随机游走将不再满足递归性。在一个房间里到处飞的苍蝇和在空气中弹跳的分子都只会有限次地返回到它们的起点。[16]（正因为如此，才会在本章开头引用角谷静夫的那段话。）

随机游走的无递归性为模型如何阐明我们的思考提供了一个很好的例子。直觉告诉我们，当添加维度时，返回起点的次数应该会减少，而逻辑则表明，这里会出现一个突然的变化。在一维和二维的情况下，随机游走会无限次地返回起点。而在三维的情况下，它将“永恒在外游荡”。要得到这种结果必须利用数学，只靠直觉是不够的。

使用随机游走估计网络规模

我们可以利用低维随机游走的递规性来估计某个网络的规模。方法很简单，随机选择一个节点，然后沿着网络的边开始随机游走，并跟踪它回到初始节点的频率。返回到初始节点所需的平均时间与网络的规模相关。例如，为了估计一个社交网络的大小，可以要求某人指定一个朋友，然后让那个朋友再说出一个朋友的名字，一直继续这个过程，看需要多久才会返回到同一个人。

图13-2显示了两个网络。左边的网络有3个节点，它们组成了一个三角形。右边的网络有6个节点，组成了两个三角形。在左边的网络上，我们不妨从A开始随机游走。假设它先移动到B，然后再移动到C，最后再返回到A。这也就是说，随机游走只需3步就可以返回它的起点。而在右边的网络上，从D开始的随机游走可能需要7步（F—G—H—F—E—F—D）才能回到起点。如果将这样的实验重复多次，那么左边网络的平均返回时间显然会比右边网络要短。虽然对这些小型网络来说，要衡量它们的规模并不一定需要这种方法，但对于大型网络（如万维网或大型电子邮件网络）来说确实非常有用。

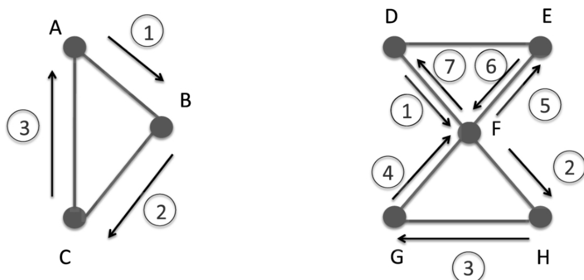


图13-2　网络上的随机游走

随机游走与有效市场

事实已经证明，股票价格接近正态随机游走，带有正漂移，以获得市场收益。许多个股的价格也接近随机。图13-3显示了Facebook在2012年5月18日首次公开发行后一年中的每日股票价格数据。Facebook公开发行时的价格为每股42美元。截至2012年6月1日，股票价格已经下跌到了28.89美元。一年后，价格进一步下降至24.63美元。图13-3还显示了另一个已经校准为具有类似变差的随机游走。

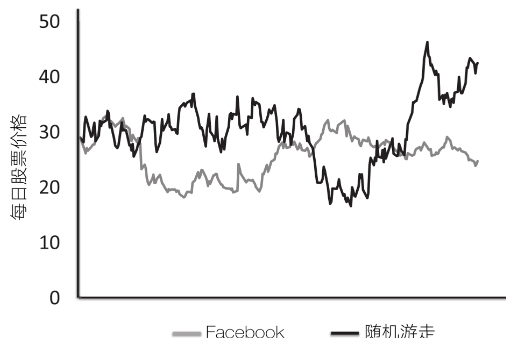


图13-3　2012年6月—2013年6月，Facebook每日股票价格vs. 一个随机游走

我们可以对Facebook的股价序列进行统计检验，以确定它是不是真的满足正态随机游走的假设。首先，价格应该以相同的概率上下波动，在这个序列所涵盖的249个交易日内，Facebook的股票价格在127天内是下跌的，占总交易日数的51%。其次，在随机游走中，增加的概率应该与前一周期的增加无关，Facebook的股票价格连续两天在同一方向上发生变化的时间只占总时间的54%。最后，持续出现在同一方向上的最长波动应该是8天，在这一年时间里，Facebook的股票价格曾连续10天上涨。因此，总的来说，我们不能否认Facebook的股票价格与正态随机游走一致的假设。

同样的分析也适用于所有股票的日交易价格。为了做到这一点，我们必须先去除股票价格中所包含的平均上涨趋势。研究表明，从20世纪50年代到80年代，每日股票价格略有正相关关系。在进行了去趋势处理之后，一天上涨之后再出现上涨的概率略超过50%。20世纪80年代之后，由于投资者开始变得更加精明，一天上涨之后再出现上涨的概率下降到了50%，从而与随机游走完全一致。

股票价格可能遵循随机游走的原因是，聪明的投资者能够识别出并消除这种模式。例如，在20世纪90年代，分析师注意到，股票价格往往会在每年年初出现上涨，这种现象被称为“一月效应”（January effect）。聪明的投资者可以在12月以低价购买股票，并在来年1月卖出以获取利润。这个策略看起来好得让人难以置信，而事实是，如果投资者在12月购买股票，他们就会抬高价格，从而抵消“一月效应”。事实上，我们不应该对“一月效应”的消失感到惊奇。

经济学家将市场价格的可识别持久模式类比为人行道上的百元钞票。如果有人看到人行道上有张一百元的钞票，就会把它捡起来，然而只要这样做了，钞票就会消失。同样的逻辑适用于股票价格模式：如果

它们存在，它们就会消失。因此，充满了聪明的投资者的市场几乎必定不会包含什么可预测的价格模式。既然价格不会呈现出任何模式，那也就只能是随机游走了（需要注意的是，必须先去除一般的上行趋势）。

经济学家保罗·萨缪尔森（Paul Samuelson）构建了一个能够生成随机游走的模型。他的模型并不要求投资者知晓未来所有期间的股票价值，而只要求他们知晓股票价值的分布。正如萨缪尔森所说："人们不能过于迷信现有的定理，它不能说明实际的竞争性市场运作良好。"[17]但不是每个经济学家都能认同他的观点。

一些经济学家将这种随机游走思想进行了扩展，提出了有效市场假说（efficient market hypothesis）。这个假说指出，在任何时候，股票的价格都反映了所有的相关信息，未来的价格必定遵循随机游走。有效市场假说依赖于一个自相矛盾的逻辑。[18]因为要确定准确的价格需要付出时间和精力，财务分析师必须收集数据并构建模型。如果价格真的是随机游走的，所有这类活动都将无法得到预期的回报。然而，如果真的没有任何人花费时间和精力去估计价格，那么价格就会变得不准确，也就意味着人行道上会铺满百元钞票。

简而言之，正如格罗斯曼和斯蒂格利茨悖论（Grossman and Stiglitz paradox）所强调的，如果投资者相信有效市场假说，他们就会停止分析，从而导致市场效率低下；而如果投资者认为市场效率低下，他们就会应用模型进行分析，从而提高市场效率。

事实上，股票市场上的价格变动与随机游走确实相当接近，尽管利用复杂的统计技术确实能够揭示某些短期模式。[19]这也就是说，虽然人行道上可能没有铺满百元钞票，但是在草地上确实能够找到一些四叶草，只要足够努力。

有些批评有效市场假说的人还指出，许多投资者持续战胜市场的时间明显不能用偶然性来解释。[20]此外，股票价格之所以随机波动，也可能是由于一些其他原因，例如复杂的交易规则的总体影响。日常价格的波动性超过了流入市场的信息量，而且，在现实世界中似乎并没有发生什么重要的事情时，股票市场也会出现大幅飙升或跳水，这就表明市场上存在泡沫。给某个人带来很大不利的某个事件，对另一个人来说也许不过是"尽管有这些问题，但是……"。是的，波动性确实很高，但是很少的信息就可能会产生很大的影响。即便市场真的出现了大幅飙升或跳水，市场仍然可能是遵循长尾随机游走的。在长尾随机游走中，股票的日常波动源于长尾分布。

虽然，股票价格始终准确的说法似乎令人难以置信，但从长远来看，价格确实不会与真实价值相差太远。我们可以应用72法则来证明这一点。如果经济每年增长3%，那么在半个世纪中，经济总量将增长4倍。如果回到1967年，当时美国的国内生产总值相当于今天的4.2万亿美元（按2009年美元计算），而到了2017年，美国的国内生产总值增长到了将近17万亿美元（按2009年美元计算），增长了4倍，这正是我们所预期的：每年增长3%，半个世纪就可以增长4倍。在同一时期，标准普尔500指数股票的实际价值也增加了大约4倍。如果股票市场每年上涨12%（以实际美元价值计算），那么股票价格就会增加256倍，这应该是不可能的。[21]

从长远来看，有效市场假说或类似的假说是合理的。但是从短期来看，押注价格修正却可能存在不小的风险。在这方面，长期资本管理公司（LTCM）的成败经历很有启发性。这是一家对冲基金，其董事会成员中包括了两位诺贝尔经济学奖得主。在1996年和1997年，长期资本管理公司公布的年回报率都超过了40%，原因是它发现了市场上的效率低下问题，并预测市场会做出修正。1998年，他们（正确地）注意到了俄罗斯债券价格与美国国债价格之间的不一致性，于是他们下了一个很大的赌注。然而，俄罗斯的违约（自1917年以来的首次违约）在短期内进一步增大了这种不一致性。长期资本管理公司一下子亏损了46亿美元，几乎导致整个金融市场崩溃。在长期资本管理公司得到救助后不久，债券价格确实恢复了一致性，但是这个修正来得太慢了。长期资本管理公司给我们的教训是深刻的，也是显而易见的，那就是不要过分相信一个模型。

小结

在本章中，我们讨论了伯努利瓮模型和随机游走模型，然后将这些模型应用到了很多领域。我们看到，这些模型能够从连续得分现象中解析出随机性，还可以用来制订博彩策略、评估股票价格变化的时间序列，以及理解篮球比赛的结果。我们还懂得了如何应用随机游走返回时间的幂律分布来增进对企业生命周期和生物分类单元的理解。

从这些应用中，我们看到随机游走模型为评估时间序列提供了一个很有用的框架。我们不能被一两年的成功所愚弄，因为那并不意味着持续的卓越。在《从优秀到卓越》（*Good to Great*）这本有史以来最畅销

的商业书籍之一中，吉姆·柯林斯（Jim Collins）描述了那些能够持续取得成功的公司的特点，例如拥有谦逊的领导者、选择合适的人进入团队、保持严格的纪律。柯林斯以6次铁人三项世界冠军戴夫·斯科特（Dave Scott）的习惯为例，将之称为“冲洗你的奶酪”。戴维·斯科特会清洗奶酪以减少身体的脂肪含量。柯林斯在这本书中特意列出了11家坚持了他所说的那些原则的“伟大公司”。但是，在他的书出版后的10年中，只有一家公司实现了强劲增长。另外10家公司中，一家被其他企业收购了，一家由政府接管了，另外8家则只带来了零回报。

伟大的企业确实会拥有一些共同特征，但这个事实并不意味着这些特征就必定有助于成功。也许，很多表现糟糕的公司也拥有这些特征。挑选一些看上去很好的公司出来，列出它们的特征，这并不是模型思维。模型思维的要求是，推导出能够导致成功的那些特征，例如才华横溢的工人，然后再根据数据来检验相关结论。如果可能的话，最好寻找一些自然实验，也就是相关特征随机变化的实例。

其他模型，例如我们在第28章中将会介绍的舞动景观模型和崎岖景观模型，更是对吉姆·柯林斯的全部理论提出了根本性的质疑。如果经济是复杂的，那么今天证明成功的特征在未来并不一定同样有效。按“大石头优先”原则，当前的“伟大”在10年后甚至有可能连“不错”都算不上。在得出一般性的结论之前，必须应用多个模型，以避免“犯大错”的风险。也应该注意避免被某些“模式”所惑，看上去似乎是一个趋势，其实可能是随机。

14　路径依赖模型

人不能两次踏进同一条河流，因为无论是这个人，还是这条河，都已经不同了。

赫拉克利特（Heraclitus）

本章将讨论路径依赖模型。在任何领域，人们的行为都建立在他人行为的基础上，无论是国际事务、艺术、音乐、体育、商业、宗教、技术还是政治，我们都应该会看到某种程度上的路径依赖。大学生选修的课程，导致他未来会倾向于选择某些职业而不会选择其他职业；对某个候选人的认可，可能会使他开启政治生涯；友谊可能会导致其他社会关系；我们穿的衣服、读的书、看的电影，以及任何需要我们付出时间的活动，都会表现出一定程度的路径依赖性。

路径依赖也存在于更大的尺度上。普通法裁决会确立并强化先例，从而影响未来的裁决。[1]早期的制度形式会影响后来的制度选择。美国政府制定的由私营公司提供健康保险的政策，导致了一个宏大的私营医疗保险行业以及一系列保健机构，还导致了公立医院和私立医院并存的结果。[2]特定的制度还会引发特定的行为模式，如自私或合作的倾向；而这种倾向反过来又会影响未来制度的效力。[3]

在本章中，我们构建了一系列动态瓮模型（dynamic urn model），该模型将生成表现出路径依赖性的结果序列。这种模型扩展了伯努利瓮模型，允许瓮内球的分布随过去的结果而变化。利用这些模型构建我们的思维，给出路径依赖的正式定义，并将路径依赖的结果与均衡区分开来。这些正式的定义还可以将路径依赖与临界点区分开来，而临界点是结果中更加突然的变化。

本章由四个部分组成。前两个部分分别讨论了波利亚过程（Polya process）和均衡过程（balancing process）。波利亚过程假设正反馈并产生路径依赖的结果和均衡。关于路径依赖的许多典型例子，包括“QWERTY”打字机的发展，都是基于正反馈的，这种情况也被称为收益递增。均衡过程则假设负反馈并产生路径依赖的结果，但不会产生路径依赖的均衡。第三部分定义了一个基于熵的路径依赖度量，最后一个部分则讨论了模型的进一步应用。

波利亚过程

波利亚过程利用伯努利瓮模型的扩展来刻画正反馈效应。在波利亚过程中，我们会往瓮中加入与抽取出来的球相匹配的球，这个过程会产生结果路径依赖（outcome path dependence）。结果路径依赖是指每一周期的结果都取决于先前的结果。说结果的长期分布，也就是均衡路径依赖（equilibrium path dependence）也取决于结果，也是对的。[4]这两种类型的路径依赖之间的区别，是本章要讨论的一个核心内容。一个均衡路径依赖的过程，必定是结果路径依赖的。如果长期的结果取决于路径，那么这个路径“沿途”的结果也必定如此。不过，一个过程是结果路径依赖的，并不意味着它一定是均衡路径依赖的。现在发生的事情可能取决于过去，但是长期均衡却可能在一开始就确定下来了。

波利亚过程

一只瓮里面装着一个白球和一个灰球。每一周期，都随机抽取出一个球并将这个球与和它颜色相同的另一个球一起放回到瓮中。抽取出来的球的颜色表示结果。

波利亚过程可以用来刻画多种多样的社会和经济现象。一个人选择学习打网球，还是打壁球，可能取决于其他人的选择。如果更多的朋友选择学习打网球，那么这个人就更有可能也选择学习打网球，因为这会增加他找到伙伴打比赛的机会。与此类似，一个人决定购买什么类型的软件、学习哪种语言或购买哪款智能手机，也可能取决于他的朋友以前做出的选择。类似的逻辑同样适用于企业对技术标准的选择，它们可能会根据其他企业的选择来做出选择。

这个模型通过改变球的分布来刻画这些社会影响。如果灰球代表选择学习打网球的人，白球代表选择学习打壁球的人，那么，当更多的人选择学习打网球时，瓮里面就包含了更多的灰球，从而导致后来的人更加有可能选择学习打网球。对更多人所选择的结果的这种不断增长的牵引力创造了路径依赖。

我们可以从波利亚过程中推导出两个令人意想不到的性质。首先，具有相同数量的白色结果的任何序列都会以相同的概率发生。其次，白球和灰球的每个分布都以相同的概率发生。第二个性质意味着极端的路径依赖。任何事情都可能发生，一切皆有可能。在1 000个周期之后，

瓮中包含了40%的白球的概率等于包含了2%的白球的概率。

要理解为什么会这样，不妨考虑前三个周期内所有可能的结果序列。第一个周期的结果是灰色的概率为1/2。如果这个结果发生了，那就要加入一个灰球，从而使第二个周期的结果为灰色的概率增加到2/3。如果第二个周期的结果也是灰色，那就要加入第三个灰球，从而使第三个周期的结果为灰色的概率增加到3/4。由此可以得出，三个灰球（或三个白球）的总概率等于1/2乘以2/3乘以3/4，其乘积等于1/4。

前三个结果由两个白球和一个灰球组成的结果序列如图14-1所示。在最上面一行中，结果的顺序是先灰色、再白色、然后又是白色。这个序列的概率是1/12，其他两个序列的概率也是一样。因此，获得这三个序列之一的概率等于1/4。根据对称性，选择两个灰球和一个白球的概率也必定等于1/4。因此，每个可能的结果集——三个白色、三个灰色、两个白色以及一个灰色、两个灰色、一个白色，出现的概率都是1/4。而且，产生两个白色和一个灰色的每一个序列都以相同的概率出现。对任意周期数，都可以得到类似的结果。[5]

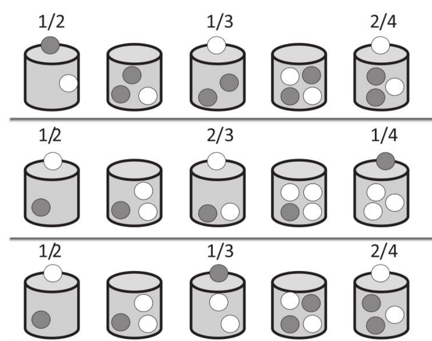


图**14-1**　由两个白球和一个灰球组成的各种结果

如果我们将波利亚过程加以扩展，也就是进一步加入其他颜色的球，这两个性质仍然成立。各种颜色的任何比例都会出现而且概率相同。这些结果为消费品生产商带来了一个难题：消费者对产品某些性质的长期偏好可能是随机的。无法预测结果的知识仍然可以用来指导行动。例如，福特公司不会希望出现这样的情况：生产出了4万辆黄色皮卡车，后来却发现在一个路径依赖的过程中，红色成了消费者最喜欢的颜色。不受欢迎的颜色的产品会大量积压在库存中，这种可能性指向两种可能的行动。生产企业可以考虑建立这样的供应链：把颜色的决定放

在最后一个环节。例如，服装公司可以先不给毛衣染色，直到看清楚流行的颜色是什么为止。或者，生产企业也可以不给消费者选择的机会。亨利·福特只提供一种颜色的T型车，即黑色T型车。苹果公司在推出第一款iPhone时也做了同样的事情：你可以买一部黑色的iPone，或者以同样的价格买另一部黑色的iPone。

均衡过程

第二个模型是均衡过程，它的假设与波利亚过程恰恰相反。在抽取出某种颜色的球后，要加入一个相反颜色的球。如果在前两个周期都抽取出了白球，那么瓮中将包含三个灰球和一个白球，从而导致下一周期抽取出灰球的概率增大为3/4。这个过程也会产生路径依赖的结果，因为任何一个周期结果的可能性取决于过去的结果的历史。但是，它不会产生依赖于路径的均衡。从长远来看，瓮收敛为每种颜色的球的比例都相同。[6]

均衡过程

一个瓮包含一个白球和一个灰球。每一周期都随机抽取出一个球，并将与抽取出来的球颜色相反的球与抽取出来的那个球一起放回到瓮中。球的颜色表示结果。

均衡过程可以用来刻画有趋向平等分配压力的决策或行动序列。有两个孩子的父母可能会尝试给每个孩子分配相同的时间。与一个孩子共度了一个下午之后，父母会产生与另一个孩子共度更多时光的愿望。

平衡过程甚至可以用来对努力实现公平的组织行为建模。国际奥委会希望世界各地都能举办奥运会。2013年，国际奥委会宣布东京成为2020年夏季奥运会和残奥会的主办城市，与东京竞争的两个欧洲城市伊斯坦布尔和马德里败北。四年后，国际奥委会决定把2024年和2028年奥运会主办权分别授予欧洲城市巴黎和北美城市洛杉矶。东京之所以能够赢得2020年奥运会的主办权，部分原因当然是因为它很有竞争力，但是还有一部分原因则是因为自1964年以来夏季奥运会从未在日本举行过，在地域上保证公平这种考量似乎发挥了作用。欧洲、亚洲和大洋洲以及美洲在第二次世界大战后举办奥运会的次数大体上相同：欧洲8次、美洲6次、亚洲和大洋洲7次。

路径依赖还是临界点

路径依赖是对结果的逐渐影响，临界点则意味着结果的突然变化。微软公司的发展为路径依赖提供了一个很好的例子。微软公司于1975年成立后，为BASIC计算机语言开发了解释器。1979年，微软与国际商用机器公司（IBM）达成协议，为IBM的个人电脑提供操作系统。这笔交易使微软走上了将一个只拥有40名员工的公司转变为世界上最有价值公司之一的道路。

IBM的合同促进了微软的发展，但是并不能保证微软长期的成功。当时，个人电脑的市场还很小，互联网并不存在，也没有复杂的文字处理程序、商业软件或视频游戏。此外，个人计算机的成功还部分取决于微软开发的DOS操作系统。在个人计算机市场发展起来后，其他公司也开发出了很多与DOS兼容的软件，从而提供了更多的正反馈。DOS的成功、个人计算机市场的发展以及与DOS兼容的软件的开发，都可以被视为从瓮中抽取出来的始终是同一种颜色的球。每个结果都使下一个结果更有可能。计算机时代的到来也许是不可避免的，但是微软的核心作用和个人计算机的发展则只代表了许多种可能的发展路径之一。

我们可以将微软公司成长的路径依赖与1914年6月28日弗朗茨·斐迪南大公（Archduke Franz Ferdinand）遇刺事件对比一下，许多人都认为，这个暗杀事件是导致第一次世界大战爆发的临界点。在暗杀发生6年前，奥匈帝国吞并了波斯尼亚和黑塞哥维那。很多塞尔维亚人都对这种情况心怀不满，于是加夫里洛·普林斯（Gavrilo Princip）枪杀了弗朗茨·斐迪南大公夫妇。奥匈帝国将这个事件归咎于塞尔维亚政府，这是一个几乎不可避免的反应。奥匈帝国准备向塞尔维亚宣战时，决定向德国靠拢，寻求德国皇帝威廉二世的支持。紧张局势迅速升级。塞尔维亚随后与俄罗斯结盟，而俄罗斯又与法国和英国结盟。到了1914年8月2日，德国就向法国宣战了。8月3日，比利时拒绝让德国借道进入法国，世界大战全面爆发。这个过于简化的事件版本表明，考虑到世界各国的结盟状况，斐迪南大公的遇刺是世界开始向战争倾斜的临界点。

我们可以利用可能结果的概率变化来度量路径依赖和临界点。[7]对于波利亚过程，初始概率在瓮中的所有分布上都是均匀的，这是一个最大熵分布。随着事件的展开，分布逐渐变窄，标志着路径依赖的形成：当结果出现后，可能发生的事情也会变化。这种熵的减少是渐进的。而对于临界点，概率分布是突然改变的，熵可能会迅速下降。图14-2显示了

这两个过程的差异，每个过程都有两个可能的结果。事件发生后，也就是微软获得了IBM的合同或斐迪南大公遇刺后，每次的概率就变了，后续事件也会改变概率。具有临界点的过程出现大幅度的转折，而路径依赖的过程则变化缓慢。

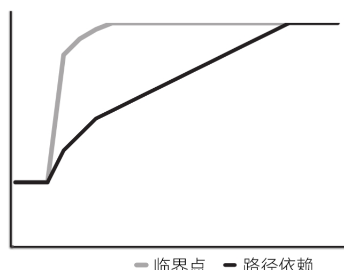


图14-2　临界点与路径依赖

路径依赖模型的应用

在现实世界中，路径依赖可能不会像波利亚过程那样极端。然而，我们还是可以从模型中推断出当行为具有很大的社会性成分时，几乎任何事情都有可能发生。在一个大学里，大多数学生在冬季都可能穿黑色大衣，而在另一个大学里，学生们则可能穿孔雀蓝大衣。模型思维告诉我们，行为差异既可能是社会影响的结果，也可能是不同的内在偏好所致。在任何时候，只要人们在一组固定的备选项中进行选择，而且他们的选择依赖于其他人先前做出的选择时，就会出现这种情况。这样的例子包括在选举中投谁的票、要看哪一部电影，以及购买哪一种技术等。

模型还可以扩展到允许社会影响因所选方案的不同而变化。香草冰激凌可能会有一定程度的反馈。而具有异国情调的绿茶冰激凌则可能得到更具多样性的反馈：一个朋友可能不喜欢它并且劝你不要尝试，另一个朋友可能很喜欢它并鼓励你一定要试一试。这可以证明，反馈的变化越少，选择这种结果的可能性就越大。[8]模型也可以改变，使人们对社会影响的敏感性不同，也就是说，人们对加入瓮中的球赋予不同的权度。

在模型的任何变体中，我们都可以测量（或估计）路径依赖的程度并将之与其他模型的变体进行比较。如果我们在构建一个关于新产品的引入模型时所做的假设表明结果取决于路径的早期部分，那么尽早进入、进行干预或补贴可能是一个不错的策略。这个模型为企业尽早将产品推向市场或提供大幅度的折扣以便尽可能多地吸引早期用户提供了逻

辑支持。当然，其他假设则可能表明，拥有更好的产品可能比更早进入市场还重要，也就是说，更好的策略是关注质量。利用各种模型，我们可以识别出与特定情况相关的不同特征，比如个人偏好与社会影响的相对重要性及反馈的变化以及质量的相对差异，并利用这方面的知识去制订策略和指导数据收集。

最后，波利亚过程表明，正反馈是怎样生成结果路径依赖和均衡路径依赖。路径依赖现象会出现在各种各样的情况下。只要一个行动会与未来行动“相遇”或与未来行动相互作用，就会出现某种程度的路径依赖（如果不是均衡路径依赖，也会是结果路径依赖）。在做出关于大型公共项目的决策时也是如此。[9]建造公园或高速公路的决定会对未来的规划决策构成约束。路径依赖程度通常取决于项目的规模。例如，纽约中央公园对纽约市的发展就产生了深远的影响。虽然波利亚过程揭示了交互导致路径依赖的这个核心思想，但是要利用这个洞见来指导行动，还需要更加真实的模型。

风险价值与波动性

我们可以将时间序列数据中的标准差解释为波动性。对股票、房地产和私人股权的投资都会表现出波动性。风险价值（value at risk，VaR）衡量的是在某个特定时间段内损失特定金额的概率。一项风险价值为一年5%的总额为1万美元的投资，在一年结束时损失超过1万美元的可能性为5%。[10]银行使用风险价值法来确定为了避免破产手头必须持有的资产数量。例如，为了保证一项风险价值为两周40%的1万美元投资的安全，投资者可能被要求必须持有10万美元的现金。

如果投资遵循简单随机游走，且每一周期投资规模的增大量或减少量为M，那么它的风险价值为N期2.5%的$2M\sqrt{N}$。[11]因此，每天随机上涨或下跌1 000美元投资的风险价值为9天2.5%的6 000美元，或者，一年2.5%的38 000美元。考虑到风险价值在步骤大小上是线性增长的，但它会像周期数的平方根一样增长，我们可以使用风险价值来解释为什么美国联邦存款保险公司（Federal Deposit Insurance Corporation）只要求银行保留相当于资产总额的2%隔夜现金，而银行则要求消费者支付房子总价的20%。这种差异是由于贷款期限不同所致。隔夜贷款的期限是一天，住房贷款可能持续数十年。3 065天的平方根约为60。

这里，我们假设了一个正态随机游走，计算风险价值的分析师通常考虑过去的经验收益分布。如果经验分布有一个较长的尾部，也就是说，如果它包含更多的大事件，那么风险价值会随着大事件的发生而增加。

虽然风险价值起源于金融领域，但是它所包含的思想却是可以广泛应用的。一个由非营利组织运营并由志愿者提供服务的施粥处通常需要25名志愿者，该组织可能想了解缺乏足够志愿者的可能性。如果志愿者的数量是一个每个星期增加或减少一个人的简单随机游走，那么就可以利用上面的风险价值公式，并且将参数设定为M＝1和N＝52，从而可以确定，风险价值为一年2.5%的15（在一年、2.5%的概率下，风险价值为15），这意味着该非营利组织有2.5%的概率会缺少一名志愿者。

15 局部互动模型

每一代人都嘲笑旧时尚，同时又虔诚地追随新时尚。

亨利·戴维·梭罗（Henry David Thoreau）

在本章中，我们要研究两个局部互动模型（local interaction model），即局部多数模型（local majority model）和生命游戏。这两个模型都是建立在一个由单元格组成的棋盘上，棋盘由处于两种状态之一的元胞组成。表面上看，这两个模型似乎没有太大的不同，其实不然。在局部多数模型中，元胞通过与它的大多数邻居的状态相匹配来更新。而在生命游戏中，元胞的更新规则要更加复杂，它依赖于多个阈值。这两个模型的结果也不同。局部多数模型总是收敛到均衡，生命游戏则取决于其初始条件，可能会产生任何类型的结果：均衡、周期性、复杂性或随机性。

局部多数模型可以用于解释和预测社会系统和物理系统中的实际结果。它可以通过符合个体或者刻画像自旋玻璃这样的磁性粒子与自己的“邻居”对齐的物理系统来表示离散选择。相比之下，生命游戏则完全是探索性的。开发生命游戏的目的就在于探索简单的规则如何集结并产生各种复杂的现象。在生命游戏中，周期性模式、复杂序列和随机性都可能从相互作用中涌现出来。生命游戏突显了整体与部分的不同。这里，我们可以给出一个粗略的类比：在人类的大脑中，诸如情绪、认知和意识等涌现现象也会从更简单的部分涌现出来。

首先分析局部多数模型。我们将阐明，一个标准的协调博弈如何为模型所假设的行为规则提供微观基础。因此，我们既可以将模型中的参与者解释为遵循规则的行为者，也可以解释为应用某个最优反应策略的理性行为者。然后，我们分析了生命游戏，并说明它如何从简单的规则中产生复杂性。本章末尾的讨论强调了利用多个局部互动模型探索的价值。

局部多数模型

局部多数模型假设元胞是排列在棋盘上的。[1]每个元胞处于两种状态中的一种：开或关。初始时，我们随机地给元胞分配状态，此后，元胞

的状态取决于它“邻居”的状态。邻居可以通过多种方式加以定义。我们将元胞C的邻居定义为位于它东、南、西、北的4个元胞以及4个对角上的相邻元胞，因此它的领域大小为8。

局部多数模型

二维方格上的每个单元都处于两种状态之一：开或关。每个单元有8个邻居。[2]在每个周期中，随机选择一个元胞。[3]当且仅当其中它的5个或更多邻居处于另一个状态时，这个元胞才会改变自己的状态。

1	2	3
4	C	5
6	7	8

局部多数模型中的局部互动包括正反馈：元胞要与其他元胞的状态相匹配。图15-1显示了局部多数模型的典型均衡配置。

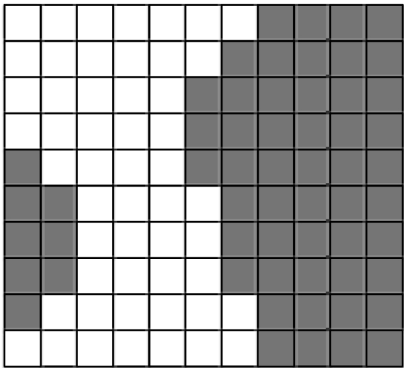


图15-1　局部多数模型中的典型均衡配置

在均衡状态下，每个元胞的状态都与它的大多数邻居的状态相匹配。均衡配置类似于荷尔斯坦奶牛身上的黑白斑块。虽然均衡配置取决于各元胞的初始配置，但是该模型对初始条件没有表现出极高的敏感性。改变一个元胞的状态最多只会导致最终配置出现很小的变化。最终的模式还取决于元胞被激活的顺序。因此，这个模型也表现出了路径依赖。它的均衡数量非常巨大。模型产生的两个均衡之间的相似度，不会比同一块地里的两头荷尔斯坦奶牛身上的黑白斑块更高。

局部多数模型最初是为了用来刻画一类物理系统而开发的。在这类

物理系统中，每个元胞的状态代表一个原子的自旋态，我们不妨将每个元胞想象为一个具有负电荷或正电荷的磁体。每个磁体都位于一个局部磁场中，物理力量会驱使这个磁体与它的邻居的自旋相匹配。同样的模型也可以用来表征玻璃体和水晶体。

在这里，我们利用这个模型来刻画人与人之间的局部协调或一致性。假设每一个元胞都代表一个人的行动。这里的行动可以是任何行为惯例，例如握手或鞠躬，打断或举手。任何一个人都想选择一个与邻居相匹配的行动。棋盘代表社交网络，在很多情况下，棋盘确实是代表社交网络的适当模型。例如，当房主决定维护自家的院子、植树造林或创建生态景观的时候，又或者礼堂内的人决定是否为表演者起立鼓掌时。[4] 虽然棋盘最多只能作为一个粗略的近似，但是我们利用它可以得出核心直觉结论。

如果在计算机上对这个模型进行模拟，就会发现它总是能实现斑块状的均衡配置。在第16章中，我们将给出这种均衡出现的原因。在局部多数模型的物理解释中，斑块状均衡模式对应于受挫状态（frustrated state）。许多元胞在开的状态下有一些邻居，在关的状态下也有一些邻居。

如果我们通过社会视角解释模型，就可以把这种受挫状态理解为次优均衡（suboptimal equilibrium）。如果开对应于以握手的方式表示欢迎、关对应于以鞠躬的方式表示欢迎，那么位于“斑块”边界上的人在与他们的邻居互动时可能会碰到一些尴尬的情况：当其他人准备握手时他们却准备鞠躬，或者当其他人准备鞠躬时他们却准备握手。如果每个人都选择了相同的动作，也就是说，如果他们解决了协调博弈问题，那么从整体上看，所有人都会更加开心。

由于这种互动只作用于局部，因此出现了次优均衡，即受挫状态。相反，如果元胞是根据全局多数原则来匹配的，那么很快所有元胞都会处于相同的状态。这种观点意味着，创建共同行为可能需要影响更加广泛的网络。如果人们只在局部与自己的邻居协调，他们就会创造出各种各样的行为。因此矛盾的是，恰恰是协调导致了多样性。

纯粹协调博弈

在纯粹协调博弈（pure coordination games）中，每个博弈参与者都要在两个行动A或B中选择一个。如果两个博弈参与者都选择了相同的行动，那么每个博弈参与者的收益为1。如果两个人各自

选择了不同的行动，那么每个人的收益都为零。

行动	A	B
A	1，1	0，0
B	0，0	1，1

一个纯粹协调博弈有两个有效的均衡：两个博弈参与者选择A或者两个博弈参与者选择B。它还有一个无效的均衡，也就是每个博弈参与者都在A和B之间进行随机选择。我们可以从纯粹协调博弈的角度重新解释局部多数模型：在这个博弈中，每个元胞都是一个博弈参与者，它必须选择一个共同行动来对抗它的8个邻居。如果博弈参与者只有在随机激活时才能改变行动，那么某个博弈参与者就可以通过选择与它的大多数邻居所选择的行动相匹配的行动来增加自己的收益。这种策略被称为短视最优响应（myopic best response），因为它没有考虑到邻居可能的未来行为。

一个博弈参与者有5个邻居选择了B，那么这个博弈参与者可以通过从A切换到B来在短期内增加自己的收益，但是，如果这个博弈参与者和邻居是被其他选择A的博弈参与者的“海洋”所包围的一个孤岛，那么这个博弈参与者保持A不变反而可能获得更高的期望收益。最关键的一点是，局部多数模型中的行为规则虽然只是一个假设的规则，但却可以植根于博弈论模型中。

协调的悖论（paradox of coordination）将不同群体之间的差异解释为一种特异性的分歧。对于某些行为，比如你是将酱油或番茄酱存放在橱柜中还是冷藏在冰箱里，或者当别人走进你家里的时候，是会脱鞋进门还是直接穿着鞋进来，而与其他人协调一致是明智的选择。由此而产生的地区多样性使我们的生活更加丰富多彩。以咖啡为例，意大利的小杯力士烈特（ristretto）、法国的中杯浓缩，以及华沙的大杯咖啡加奶油（kawa ze smietanka）极大地增加了人们在欧洲旅行的乐趣。

但是，其他很多差异却可能是低效的。各种各样的电插头可能会令人抓狂——这里是两个插脚的，到那里却是三个插脚的。随着全球一体化程度的提高，技术协调失败的代价可能是非常高昂的。例如，瑞典政府决定将靠左行驶改为靠右行驶，以便与欧洲大陆其他地区保持一致。这件事非同小可，转为靠右行驶的那一天在瑞典被称为“H日”（Dagen H）。1967年9月3日凌晨4点45分，瑞典行驶在路上的所有车辆以及许

多特意在凌晨时分上路参加这个活动的瑞典人都突然停下来，然后，在接下来的15分钟内，所有的汽车都从左侧行驶到右侧。凌晨5点，所有汽车开始在道路的另一侧再次启动，开始靠右行驶。尽管达成协调的动机相当强烈，但是许多时候人们却无法做到。例如，虽然英国已经通过隧道连接到了欧洲大陆，但是英国的车辆至今仍然靠道路的“错误”一侧行驶，而且英国的某些（尽管不是全部）前殖民地也是如此。

协调的悖论

> 如果人们是在局部进行协调的，那么从全局的角度来看，整体配置将会是斑块状的、多样性的。

应用这个模型时，必须记住，许多帮助我们实现协调的文化习俗，例如人们如何哀悼死者的逝去、如何庆祝孩子的出生，并不是简单的“奇风异俗”，而是文化的重要组成部分，是一系列内在一致的行为规则、惯例的组合；它们定义了一个人到底是谁，给了人们生活的意义和归属感。[5]

与任何模型一样，对于局部互动模型，我们也可以对各种参数进行模拟实验，看看改变它们会如何影响结果。在局部互动模型中，均衡时形成的斑块大小增加得比邻域的大小更快。如果我们使邻域，也就是影响网格上一个格子的格子数量增加到原来的两倍，那么斑块会变得比原来的两倍还大。因此，这个模型表明，当技术和城市化使人与人之间的联系更加紧密之后，协调的力量可以产生更大的同质性行为和信念。

模拟实验还表明，如果我们将整个配置变成一个狭长的矩形，那么模型一般来说会产生水平和垂直的条纹，如图15-2所示。[6]斑马状的条纹是一个均衡，因为每个处于开（关）状态的元胞，都有5个处于开（关）状态的邻居。在正方形上，这种类型的模式也是一个均衡，但是很少会真的出现。尽管这种令人困惑的发现有可能导致我们误入歧途，陷入没有什么实证意义或理论价值的问题，但是它们也有可能引导我们得出一些深刻的洞见，从而带来更深层次的意外发现。

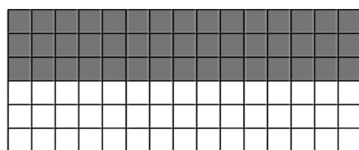


图15-2　局部多数模型中稳定的线条

在这个例子中，正方形网格会产生荷尔斯坦奶牛式斑块状模式，而狭长的矩形网格则会产生斑马状的模式这个结果，几乎就是在直接要求我们回答：是不是可以用这里的模型去解释不同动物皮毛上斑点和条纹的特有模式？大量科学研究文献告诉我们，确实可以。[7]

生命游戏

接下来我们讨论下一个模型：生命游戏。这个模型也假设位于棋盘上的元胞处于两种状态中的某一种。生命游戏与局部多数模型的关键区别在于，在这里，元胞的更新规则有两个阈值，并且所有的元胞都同步更新自己的状态。因此，在生命游戏中，我们可以说初始配置如何如何，时间1的配置如何如何，时间2的配置如何如何，等等。还可以把同步更新视为一个“进行曲”动态机制（更新！更新！更新！）。[8]

生命游戏

方格上的每个元胞都或者是活的（开的）或者是死的（关的）。每个元胞的邻居由网格上的8个相邻元胞组成。元胞根据如下两个规则同步更新自己的状态：

活的规则：对于一个死元胞，当恰好有三个活的邻居时，这个死元胞就会变活。

死的规则：对一个活元胞，当活的邻居小于两个时或当有三个以上的活邻居死去时，这个活元胞就会死去。

我们从排在同一条水平线上的三个活的元胞开始讨论，如图15-3所示。在下一个时期，对每个元胞应用上述死活规则之后，我们会得到排在同一条垂直线上的三个元胞。中间的活元胞有两个活着的邻居，所以它还活着。两端的两个活元胞分别只有一个活的邻居，所以它们都死了。最后，中间的活元胞的上方和下方的元胞都变活了，因为这两个元胞分别各有三个活着的邻居。根据对称性，等再下一个时期更新后，又会回到三个元胞组成水平线的情形。如果继续迭代运用上述规则，模型就会在水平线与垂直线之间不断交替，这也就是说，它将会不断“闪烁”。

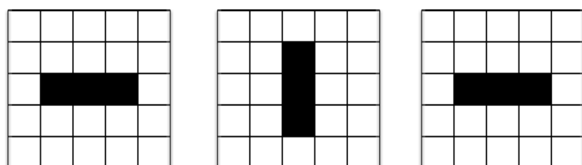


图15-3　生命游戏中的“闪光灯”

闪光灯是由元胞之间的互动产生的，而不能直接从假设中推导出来。复杂系统学者将这种宏观现象称为涌现。闪光灯是生命游戏产生的涌现结构中最常见的、最不能给人留下深刻印象的一种。图15-4显示了其他三种简单涌现结构：方块、滑翔机和右五连（R-pentomino）。这是一个均衡配置；每个活元胞都有三个活的邻居，每个死元胞最多有两个活的邻居，因此，既没有活元胞死亡，也没有死元胞复活。

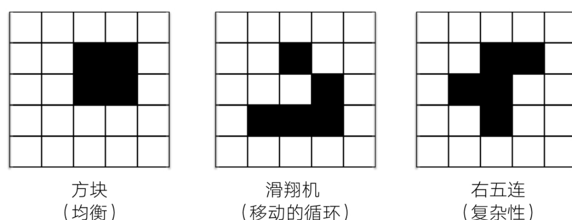


图15-4　生命游戏中的三种涌现结构

图15-4中间的配置会产生一个大小为4的循环，该循环沿着对角线向下和向右滑动，成了滑翔机。与之相关的另一个更精致的配置被称为“滑翔机枪”，它能产生无穷无尽的滑翔机。第三种配置右五连能产生一系列复杂的模式。如果我们在大型网格上重复运行模型超过1　000次，这个右五连会生成许多滑翔机、闪光灯以及其他一些较小的稳定配置。生命游戏也可以产生随机性。[9]因此，生命游戏可以根据初始状态产生任何类别的结果。

这种能力甚至引发了一些哲学问题。生命游戏由排列在网格上的、只有两种状态的元胞组成，并运用简单的规则进行更新。它可以生成精美的图案，通过适当的编码，还可以变成一台通用计算机。我们可以把初始模式视为输入，根据规则产生的结果则可以解释为计算。因此，我们可以在这个模型和人类的大脑之间进行一个粗略的类比，因为人类的大脑也是由依赖于基于阈值规则的、在空间上相互连通的简单组成部分构成的，尽管那些规则要复杂得多。

当然，这并不是说我们在生命游戏中观察到模式就可以解释意识了。现在还没有任何一本书名为《生命游戏：意识的解释》，尽管丹尼

尔·丹尼特（Daniel C. Dennett）[(6)]确实写过一本以《意识的解释》（*Consciousness Explained*）为名的书，而且他也认为像生命游戏这样的简单模型可以提供关于意识如何演化的洞见。丹尼特这个思想得到了不少人的共鸣，物理学家斯蒂芬·霍金（Stephen Hawking）就这样写道："我们完全可以想象，像生命游戏这样的东西，只有少数几个基本定律，就可以产生高度复杂的特征，甚至可能产生智能。"[10]

小结

在本章中，我们研究了两个关于排列在网格上的元胞之间的互动模型。第一个模型是局部多数模型，它总是可以达到某个均衡（尽管可能的均衡有许多种），我们可以将这个模型与多种物理过程和社会过程进行类比。第二个模型是生命游戏，它能够产生从均衡到随机的任何一种类型的结果。这个模型与现实世界之间不存在任何明确的联系。生命游戏给出了一个很好的例子，它说明，构建替代现实的模型是怎样帮助我们产生洞察力的，也就是从微观的规则中涌现出动态的宏观层面结构，这可以极大地加深我们对世界的理解。正如生命游戏所呈现的那样，整体可以执行远远超出其各个组成部分的功能。例如，如果我们把两个3×3的方格的角连接起来，画出一个斜8字图形，那么生命游戏就会产生一个长度为8的循环模式。它循环通过一系列模式，然后在恰好8个步骤内回到那个8字图形上。这个模式，从图形上看像8字，它的行为也"好像"知道它要数到8。这确实是非常惊人的。

为什么生命游戏会产生复杂性，而局部多数模型则不可避免地走向均衡？为了理解个中原因，我们还需要更多的分析工具和框架。在第16章中，我们介绍了李雅普诺夫函数（Lyapunov function），它运用差分方程对世界状态进行分类。细致地构建出一个李雅普诺夫函数之后，我们就可以解释为什么局部多数模型必定会达到平衡，而生命游戏则不一定。

最后要强调的是，正是在我们对模型进行探索的过程中，才使模型（以及现实世界）是否会产生均衡、周期性、复杂性或随机性这个问题突显了出来。确实，模型既能够回答问题，也能够提出问题。它们关上了一些门，同时又打开了更多的门。

16　李雅普诺夫函数与均衡

数学之美，只会呈现给那些更有耐心的追求者。

玛丽安·米尔札哈尼（Maryam Mirzakhani）

在本章中，我们学习李雅普诺夫函数，它给出了模型能够实现均衡的条件。李雅普诺夫函数是在按时间索引的配置系统上定义的实值函数。在每个时间步骤中，李雅普诺夫函数都要给配置分配一个值。如果配置发生了变化，也就是说，如果模型不处于均衡状态，那么李雅普诺夫函数的值就会减小一个固定量。李雅普诺夫函数也具有最小值，这意味着到了某个时间点上，它的值最终必定会停止递减。当发生这种情况时，模型就达到均衡了。例如，我们可以使用李雅普诺夫函数来证明，为什么局部多数模型一定会收敛。

本章的核心结论是，如果能够为模型构建一个李雅普诺夫函数，那么模型必定会达到均衡。在这种情况下，我们不可能得出周期性、随机性或复杂性。更重要的是我们甚至可以确定模型收敛到均衡所需的时间，这一点在为局部多数模型构造李雅普诺夫函数中可以看得很清楚。

本章的内容分为6个部分。我们首先定义了李雅普诺夫函数，然后将李普雅诺夫函数应用于对逐底竞争博弈（Race to the Bottom Game）的分析。然后，我们分别为局部多数模型和自组织活动模型构建了李雅普诺夫函数。接着，我们证明了为什么可以为某些交易市场构建李雅普诺夫函数，却不能为另外一些交易市场构建李普雅诺夫函数，同时也分析了为什么生命游戏没有对应的李雅普诺夫函数。接着，我们讨论了一个看似令人烦恼的数学问题：在不能找到李雅普诺夫函数的情况下达到均衡。本章最后讨论了均衡是否可取的问题。

李雅普诺夫函数

一个离散动态系统（discrete dynamical system）由可能的配置空间，以及一个转移规则（transition rule）组成。我们可以把配置空间视为世界的多维状态，例如生命游戏中的活元胞和死元胞的初始集合，而转移规则则是将时间t时的配置映射到时间$t+1$时的配置上。

一个李雅普诺夫函数是从配置到实数的一个映射，它满足两个假设：第一，如果转移函数不处于均衡状态，则李雅普诺夫函数的值就会减少某个固定的数量（对此，稍后会给出更多的解释）；第二，李雅普诺夫函数具有最小值。如果这两个假设成立，那么该动态系统必定达到均衡。

李雅普诺夫函数

给定一个离散时间动态系统，它的转移规则由$x_{t+1}=G(x_t)$组成。对于实值函数$F(x_t)$，如果对于所有的x_t，都有$F(x_t)\geq M$，而且存在一个$A>0$，那么下式成立，这个实值函数$F(x_t)$是李雅普诺夫函数：

$$F(x_{t+1})\leq F(x_t)-A\text{，如果}G(x_t)\neq x_t$$

如果对于G，F是一个李雅普诺夫函数，那么从任何x_0开始，必定存在一个t^*，使$G(x_{t^*})=x_{t^*}$，即该系统会在有限时间内达到均衡。

首先为逐底竞争博弈构建一个李雅普诺夫函数。这个函数刻画了该博弈的博弈参与者的策略环境，以便每一个博弈者都更愿意提供恰好低于平均水平的水平。

逐底竞争博弈

有N个博弈参与者，每个博弈参与者在每个时期都要提出一个支持水平，其取值范围为{0，1，…，100}。提出了最接近平均支持水平2/3的博弈参与者可以获得那个期间的奖励。

这个博弈可以用来解释美国各州政府削减社会项目支出的行为，例如减少对贫困人口的援助。没有任何一个州的州长或州立法机关希望公众认为自己冷酷无情，但是，又没有人愿意提供慷慨的援助贫困人口的计划，因为那会把邻州的穷人吸引过来。于是每个州都会提供一些资金，但只愿意提供低于平均水平的资金。在通过环境保护法规或制定税率时，相互竞争的国家也会有类似的动机。各国宁愿采取只对环境施加较少限制的政策，并且为了吸引企业投资，税率低于平均水平。

逐底竞争博弈是否能达到均衡，取决于博弈参与者的行为规则。如果博弈参与者随机选择支持水平，那么结果就将会是随机的。然而，考

虑到这个博弈的收益结构，随机选择支持水平并没有意义。在这里，我们假设如下行为规则，这种规则与实验研究的结果一致。[1]在第一个时期，我们假设每个博弈参与者会随机选择一个低于50的支持水平。在以后各期中，每个博弈参与者选择的水平至少等于1且低于上一期平均水平的2/3。如果上一期的平均水平已经小于零，那么每个博弈参与者都会选择零。

很容易证明，来自任何博弈参与者的最大支持水平满足李雅普诺夫函数的条件，即最大支持水平的最小值为零。而且，在每个时期内，如果支持水平采用整数值，那么最大支持水平至少会下降1。因此，到了某个时间点上，每个博弈参与者都会提议零支持水平。也就是说，博弈参与者在逐底竞争中都碰到了“底”。在这个例子中，模型产生了一个不好的结果。为了防止出现这种逐底竞争，必须改变博弈结构。例如，为了增加对贫困人口的扶持，法律可以规定由联邦政府来提供资金，或者对各州规定一个最低“扶贫”支出水平。[2]

另外，假设我们允许博弈参与者在0至100的区间内选择任何实数值，而不是整数值。再假设在每一轮中，博弈参与者选择的支持水平等于上一轮平均水平的2/3，那么平均支持水平将会随着时间的推移而逐渐降低，但是永远不会达到零的均衡。正如色诺悖论（Xeno's paradox）一样，这个过程会越来越接近于零，但是却永远不会达到零。因此，为了确保均衡的实现，还必须假设一个最小减量（A）。

局部多数模型

现在，用李普雅诺夫函数分析一下局部多数模型。我们将局部多数模型中的李雅普诺夫函数定义为总体中的总不一致性（total disagreement），即与相反状态的元胞相邻的所有元胞数量的总和。为了证明这个模型必定会达到均衡，我们必须先证明，如果一个元胞改变了自己的状态，那么总不一致性至少会下降一个固定的数量。

这个证明的数学过程并不太复杂。首先，如果一个元胞改变了自己的状态，那么相对于它的邻居而言它必定是少数。由此我们知道，那时它至少有5个邻居处于相反状态，最多有3个邻居处于相同状态。因此，当这个元胞切换状态后，与这个元胞不一致的元胞数量至少减少了两个（图16-1）。为了计算总不一致性的变化，还必须将与这个元胞相邻的元胞对总不一致性变化的贡献考虑进去。由此，那5个或更多的现在状

态一致的元胞的不一致性减少了（每个元胞减少1），同时之前状态一致的那3个或更少的元胞现在则具有更高的不一致性了（每个元胞增加1）。因此，所有相邻元胞的总不一致性至少下降了4。

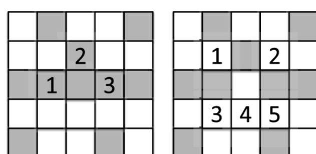


图**16-1**　局部多数模型中总不一致性的减少

这样一来，我们就证明了，即使某些元胞可能带来更多的不一致性，总不一致性也满足李雅普诺夫函数的条件。因此，局部多数模型必定会收敛到均衡，不仅仅是有些时候或大部分时候，而是所有时间都是如此。我们还可以确定收敛的速度：无论什么时候一个元胞改变自身的状态，总不一致性至少会下降4。由此可知，一个总一致性为100的配置，必定会在25个时期内达到均衡。换句话说，一个总一致性为D的配置，必定会在D/4个时期内达到均衡。正如我们在第15章中已经指出过的，实现的均衡几乎总是一个低效率的斑块状模式，它包括了不少受挫的元胞。

自组织活动模型

接下来，讨论李雅普诺夫函数的下一个应用。这一次，我们要证明自组织活动模型中也存在均衡。这个模型中由一群人和每个人可以选择去做的活动的集合组成。这个模型的关键假设是，每个人都更偏好不那么拥挤的活动，例如，更少人参加同一项活动意味着在健身房不用等待、在面包店和咖啡店不用排队。

这个模型的灵感来自著名经济学家托马斯·谢林（Thomas Schelling）的著作《微观动机与宏观行为》（*Micromotives and Macrobehavior*）。谢林在这本书中描述了城市中出现的各种令人叹为观止的自组织现象：在几乎完全不存在中央计划者的情况下，交通模式、行人流量、停留在公园和餐馆的人数，以及商店库存，到底是怎样自动达到各自的适当水平的？街头那家小店，是怎么每天都能进到4瓶来自密歇根州锡达维尔原产的纯枫糖浆的？巷尾那家面包房，又是怎么做到每天都在关门前大约20分钟卖完当天制作的黑麦面包的？尽管城市内有如此多样性的行动者——游客、店主、居民和送货人等，而且他们对整

个城市的信息所知有限，但这种秩序也会涌现出来。

自组织活动模型

一个城市里，有A种活动可以参加，每一天都由L个时间段组成。在人口规模为M的城市中，每个人都要选定一个日程安排。在这里，日程安排是指这个人在L个时间段内分配L种活动（从一个更大的K种可能性的集合中）。一个人要面对的拥挤水平则设定为等于同时选择同一种活动的其他人的数量。

为了证明这个模型是收敛的，我们要证明总拥挤度（total congestion），也就是整个人群拥挤水平的总和满足李雅普诺夫函数的条件。当一个人降低了他的拥挤水平时，就会降低自己对总拥挤度的贡献，并且会使他不再遇到的每个人的拥挤水平减少1，同时使他新遇到的每一个人的拥挤水平增加1。既然他降低了自己的拥挤水平，第一组人的数量会比第二组人更多。例如，假设一个人原本在早上8点去一个拥挤的健身房，在下午4点去一个拥挤的咖啡馆，如果她改为在上午8点去咖啡馆，并在下午4点去健身房，结果发现在上午8点那个时段咖啡馆几乎是全空的，健身房在下午4点时只是有一点点拥挤，那么她降低了自己以及之前遇到的所有人的拥挤水平。她确实使她现在遇到的少数人提高了拥挤水平，但是总拥挤度却下降了（而且至少减少了1）。既然总拥挤度不可能低于零，系统必定会达到均衡。

虽然，一般来说，我们无法保证系统能够找到一个有效率的均衡，但是这种自组织活动模型几乎总能收敛到总拥挤度近乎最小的配置。而在效率低下的配置中，更多人选择在同一个时间段参加某一项活动（比如去健身房），而不是去参加另一项活动（比如去咖啡馆）。如果在这两项活动中拥挤程度的差异很大，那么一个人就可以通过切换自己去健身房和咖啡馆的时间来降低拥挤水平。如果咖啡馆和健身房在另一个时间段有相同数量的人前往，这种切换就可以减少总拥挤度。[3]

这个模型可以解释当今世界的很多秩序，它可以让我们更加深刻地理解城市如何在没有中央计划者的情况下通过自组织实现近乎完全有效的配置。它还告诉我们，为什么许多游乐园，比如迪斯尼，却做不到这一点。迪斯尼每天都有新的入园参观者，他们没有时间去尝试新的路线。如果没有“中央计划者”的帮助，迪斯尼将会出现某些景点大排长龙，同时其他景点门可罗雀的情况。因此，为了减少这种低效率状况的

出现，迪斯尼对于某些特定的景点，只允许游人提前预约在特定时间内进入，同时让员工引导人们参观不那么拥挤的其他景点。

纯交换经济

我们也可以使用李雅普诺夫函数来探索纯交换经济什么时候可以达到均衡、什么情况下可能无法实现均衡。纯交换经济由一个消费者集合组成，每个消费者都有自己的商品禀赋和偏好。对此，我们可以设想，一群人在市场或集市上与他人交易一些东西，比如茄子、奶酪或地毯。每笔交易都需要双方付出一定的时间和精力。为了让双方有动力完成交易，每一方都必须有所获益，得到超过此交易成本的某个金额。

不过在这里，我们要做的不是直接构建一个总是按固定数量减少并具有最小值的李雅普诺夫函数。恰恰相反，我们要先证明，总幸福感（total happiness）总是按固定数量增加并具有最大值。根据假设，任何时候，只要双方完成交易，他们的幸福水平至少会提升与交易成本相当的数量。此外，由于每个人都拥有固定的商品禀赋，因此总幸福感有一个最大值。李雅普诺夫函数的假设得到了满足，系统可以达到均衡。但是，在这种均衡下，配置不一定是有效的。当然，既然不是有效的，那么市场上的一些人应该能够识别出一些让他们自己更幸福的交易。

在构建这一论点时，我们假设只有参与交易的人才有可能获得幸福（或不幸福）。但是在其他交易形式中，情况可能并非如此。试想一下，假设伊拉克与巴基斯坦达成协议，用自己的石油去交换巴基斯坦的核武器，那么这两个国家的领导人可能会因此而觉得更加幸福，但是从全球的角度来衡量，总幸福感肯定会下降，世界上其他国家可能会对伊拉克建立核武器库深感不安。

其他国家的人所感受到的这种影响被称为负外部性（negative externality）。当市场上的交易包含了负外部性时，交易就不一定能提高总幸福感了。在前面给出的那个纯交换经济的例子中，当人们交易水果、蔬菜、地毯和工具时，几乎完全不存在外部性。外部性的存在意味着我们不能直接说系统是不是会达到均衡。上面说的那种石油换核武器的交易可能会引发其他交易。例如伊拉克不断增加核武器储备，很可能导致沙特阿拉伯要求自己的盟国提供更多的军事支持，而这反过来又可能导致该地区的其他国家采取行动。在这类交易过程中，全球总幸福感或全球总安全性可能会随着每一个行动而波动不休，甚至不能确定这类

交易会不会在某一点上停止。

李雅普诺夫函数在交易环境中是否存在，取决于负外部性的大小。我以前教过的一个本科生告诉我的一个例子，可以很好地说明这一点。她所服务的公司要搬进新的办公场所，其中有一个很大的房间，配有开放式办公桌，供分析师使用。她的经理提议给所有分析师随机分配办公桌，随后让他们自由进行“办公桌交易”。经理认为这可以导致一个很好的结果，因为他认为自由交易能够带来效率的提高。

这个学生却意识到，即便任何两个相互交易的人都变得更幸福了，他们的前邻居和新邻居也不一定会有同样的感受。如果一个人决定搬到房间的另一侧，那么他当前的邻居，特别是当这个邻居是特意选择在他旁边的时候，可能会感觉受到了伤害。而且，这位前邻居也许不喜欢新邻居，因为新邻居可能会在讲电话时大喊大叫。因此，这位前邻居也可能会选择搬工位。这种搬来搬去可能会持续很长时间，每个人的士气都可能因此而受到打击。这个计划看起来相当危险。组织需要自己的员工彼此信任并相互尊重，但是在通过办公桌交易远离对方的两个人身上，是很难维持这种关系的。当经理理解了这个模型后，就放弃了他的计划。[4]

然而，故事并没有就此结束。这位经理还购买了各种风格和颜色的办公椅，计划随机分配椅子，让人们进行交易。在这种情况下，我的学生再一次使用模型思维告诉他，进行椅子交易不会产生负外部性，而且可以给员工带来不少乐趣。椅子交易是一个纯交换市场。这两种情况，对我们如何运用模型指导行动很有启发。纯交换市场适用于椅子，但是却不适用于办公桌。

没有李雅普诺夫函数的模型

有的时候，我们尝试为模型构造李雅普诺夫函数，但却无法取得成功，尽管如此，我们仍然可以积累知识。通常，我们至少能够了解模型为什么不会产生均衡。在生命游戏中，某些配置会产生均衡，而其他配置则不能产生均衡。当某个配置确实产生了均衡时，就可以写出一个特定于该配置的李雅普诺夫函数。例如，在生命游戏中，采用对角线形式的任何初始配置，每个周期长度都会减少2，因为位于对角线两端的那两个活元胞会死去，并且没有任何新的元胞会变活。这个配置以所有元胞都死光而结束。对于这样的配置，活元胞的数量就是一个李雅普诺夫

函数。如果从另一个配置开始，例如，能够产生复杂配置序列的“右五连”，就无法构造出李雅普诺夫函数来，因为系统不会达到平衡。

我们无法构建李雅普诺夫函数，并不意味着模型或系统肯定不能达到均衡。一些系统在所有已知情况下都能够达到均衡，但是没有人能够构造出李雅普诺夫函数。一个著名的例子是所谓的“取一半或乘三加一法则”（HOTPO），它也被称为考拉兹猜想（Collatz conjecture），看似简单，其实不然。“取一半或乘三加一法则”以整数开始，如果是奇数，就乘以3并加上1；如果是偶数，就除以2；当得到的结果等于1时，这个过程就停止。假设从数字5开始，由于这是一个奇数，所以将它乘以3并加1，得到16，然后，将16除以2得到8，8再除以2得到4，然后4再除以2得到2，最后得到1，达到了均衡。对于任何一个不到2^{64}的数字，“取一半或乘三加一法则”都会止步于均衡。

然而，没有人能够证明“取一半或乘三加一法则”是不是最后总能达到均衡。坊间传闻，匈牙利数学家保罗·厄多斯（Paul Erdos）曾说：“对于这样的问题，数学还不够发达。”[5]数学家们无法确定“取一半或乘三加一法则”是否总能达到均衡，这说明了一个更普遍的真理：模型所能提供的，只是证明结果的可能性。我们无法保证肯定可以推导出它们，很多时候，我们提出了一个模型之后，最终却发现要证明结果非常困难（如果不是不可能的话）。

小结

在本章中，我们已经看到，李雅普诺夫函数不仅可以帮助我们证明一个系统或模型能不能达到均衡，还可以告诉我们达到均衡的速度有多快。即便构建李雅普诺夫函数的努力遭到了失败，这种尝试也是有意义的。它们可以提供一些关于复杂性成因的线索。具有外部性的纯交换经济以及所举的交易办公桌的例子都属于这种情况。在这种情况下，我们不能构建一个总是减少或总是增加的全局变量，因此，无法保证这些过程能够达到均衡。

回顾一下模型的7大用途——推理、解释、设计、沟通、行动、预测和探索，我们就会发现李雅普诺夫函数在每个用途上都有所帮助。如上所述，利用李雅普诺夫函数，我们可以推断出系统走向均衡的原因，还可以解释系统收敛到均衡的速度、设计信息系统（例如迪斯尼世界所采取的分时段游览的预约系统）、采取行动（例如对办公桌进行交易是

不可取的）、沟通系统如何达到均衡的途径、预测系统达到平衡的时间以及进行探索。我们可以尝试提出假设和构建模型来解释一些令人惊讶的现象，例如自组织的城市。

17 马尔可夫模型

历史是一部循环诗，由时间写在人的记忆上。

珀西·比希·雪莱（Percy Bysshe Shelley）

马尔可夫模型用来刻画以一定概率在一组有限的状态之间不断转换的系统。政治体系可能在民主制度与独裁制度之间转变，市场可能在不稳定与稳定之间变化，一个人也可能在快乐、沉思、焦虑和悲伤之间转换。在马尔可夫模型中，状态之间转移发生的概率是固定的。一个国家在某一年从专制转向民主的概率可能是5%，一个人在一小时内从焦虑过渡到倦怠的概率可能是20%。此外，如果系统可以通过一系列过渡从任何一个状态转换为任何其他状态，并且不存在简单的循环，那么马尔可夫模型就可以达到唯一的统计均衡（statistical equilibrium）。

在统计均衡中，单个实体可以继续在各种状态之间移动，但是各种状态之间的概率分布仍然是固定的。例如，一个关于意识形态的马尔可夫模型，统计均衡允许人们在持自由主义立场、保守主义立场和中立立场之间转换，但是秉持每种意识形态的人口比例将保持不变。在应用于单个实体时，统计均衡意味着实体处于每种状态的长期概率不会改变。例如，当一个人处于统计均衡状态时，他在60%的时间内感到高兴，而在40%的时间里感到悲伤。这个人的精神状态可能每小时都会发生变化，但是他在这些精神状态之间的长期分布却不会发生变化。

这种独特的统计均衡还意味着，结果的长期分布不可能取决于初始状态或事件的路径。换句话说，初始条件是无关紧要的，历史也是无关紧要的，会改变状态的干预措施也不重要。随着时间的推移，满足这些假设的过程就会“不可抗拒地”走向那一个独特的统计均衡，然后保持不变。在这里，模型再一次揭示了条件逻辑：如果世界符合马尔可夫模型的假设，那么从长远来看历史并不重要。当然，马尔可夫模型并没有说历史永远是不重要的。首先，马尔可夫模型允许结果是路径依赖的，接下来会发生的事情将取决于当前的状态。其次，马尔可夫模型还允许对历史记录进行长期建模，但对于这种情况，模型的某个假设必定会被违背。

马尔可夫模型可以应用于很多领域。我们可以用马尔可夫模型来解释动态现象，例如民主转型、战争升级和药物干预，还可以用于对网

页、学术期刊和体育运动队进行排名，甚至可以用来辨别书籍和文章的作者身份。本章将介绍其中一些应用。我们将从两个简要的例子开始讨论，接着描述证明统计均衡存在性的一般定理。然后，我们转而讨论马尔可夫模型的应用。在本章末尾，会根据我们所掌握的关于马尔可夫模型的知识，重新阐述历史是如何及何时重要的问题。

马尔可夫模型的两个应用

马尔可夫模型由一组状态与这些状态之间的转移概率构成。在第一个例子中，我们将某个人在某一天的精神状态描述为“充实”或“无聊”。这两种精神状态就是模型的两种状态。转移概率表征在状态之间变动的概率。我们可以假设，当精神上“充实”时，一个人有90%的机会继续停留在该状态上，同时有10%的机会变得“无聊”；而当“无聊”时，这个人有70%的机会继续觉得“无聊”，同时有30%的概率变成一个精神上的“充实”的人。

假设上面这些转移概率适用于100名修读生物学课程的学生。在这些学生中，第一天有一半学生觉得这门课程令他们很“充实”，另一半学生则觉得很“无聊”，如图17-1所示。应用上述转移概率，可以预计第二天，会有5名原本觉得“充实”的学生（10%）会变得“无聊”，同时会有15名原本觉得“无聊”的学生（30%）会变得“充实”。这样就会有60名觉得“充实”的学生和40名觉得“无聊”的学生。到第三天，这60名原本觉得“充实”的学生中应该会有6名感到“无聊”，同时应该会有12名原本觉得“无聊”的学生变成觉得“充实”，从而导致66名学生觉得“充实”、34名学生觉得“无聊”。继续运用这个转移规则，这个马尔可夫过程会收敛到75名学生觉得“充实”、25名学生觉得“无聊”的统计均衡。在这个统计均衡中，学生们会继续在这两种状态之间转移，但是处于每一种状态的学生的总数将保持不变。

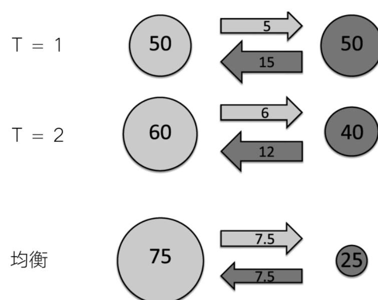


图17-1　一个马尔可夫过程

与此不同，假设这个过程开始时，初始状态是所有100名学生都觉得“充实”，那么到了第二天，将只有90名学生仍然觉得“充实”。到第三天，觉得充实的学生会下降到84名。继续迭代这个过程可以发现，从长远来看，最终仍然会收敛到75名学生觉得“充实”、25名学生觉得“无聊”的统计均衡。这个模型得出了同样的统计均衡。

接下来讨论第二个例子。在这个例子中，我们将不同国家分为三个类别：自由的、部分自由的，以及不自由的。图17-2显示了当今世界上每个类别的国家的百分比，这个分类结果是根据美国自由之家（Freedom House）截止于2010年的数据（样本期间为35年）得出的。这幅图表示，民主化呈现出了明显的加速趋势。在过去的35年间，自由国家的比例上升了20%。如果这种线性趋势一直持续下去，那么到2040年，所有国家中将会有2/3以上是自由的，而到2080年，将会有8/9的国家是自由的。

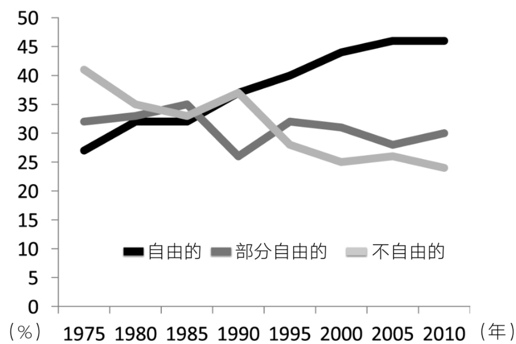


图17-2　对全世界国家的分类：自由的、部分自由的和不自由的国家的百分比

马尔可夫模型会导致不同的预测。为了进行预测，我们将每一个周期的长度设定为5年，并根据过去的数据粗略地校准了转移概率（表17-1）。[1]

表17-1　民主化的马尔可夫模型转移概率

		下一周期的状态		
		自由的	部分自由的	不自由的
当前状态	自由的	95%	5%	0%
	部分自由的	10%	80%	10%
	不自由的	5%	15%	80%

如果使用1975年时属于每个类别的国家的百分比来初始化模型，那么如我们所料，校准后的模型几乎完美地匹配了2010年的实际分布：48%的国家是自由的，31%的国家是部分自由的，21%的国家是不自由的（在2010年，这三个类别的国家的实际百分比分别为46%、30%和24%）。如果继续运行这个模型，那么它预测2080年时的情况将会是这样：62.5%的国家是自由的，25%的国家是部分自由的，12.5%的国家是不自由的。

这个马尔可夫模型的预测似乎并不乐观，原因在于这样一个事实：线性投影假设没有考虑到自由国家既可以转变为部分自由的国家，也可以转变为不自由的国家。随着越来越多的国家变成了自由的国家，从自由的国家转变为不自由的国家的数量也在增加。在现实世界中，出现这种情况的原因是多方面的。首先，实现民主要求国家财政权力机构和行政机构有很高的执行能力。借用政治学家托马斯·弗洛雷斯（Thomas Flores）和伊尔凡·努尔丁（Irfan　Nooruddin）的话来说，在某些国家或地区，民主可能不容易扎下根来。[2]在那些地方，我们有理由预期，会出现从自由国家转变为部分自由的国家的情况，马尔可夫模型也刻画了这种情况。

佩龙-弗罗宾尼斯定理

上面的这两个例子都会收敛到一个唯一的统计均衡，这并非偶然。任何一个马尔可夫模型，只要状态集是有限的、不同状态之间的转移概率是固定的、在一系列转移后能够从任何一个状态变换为任何其他状态，而且状态之间不存在固定的循环，就必定会收敛到唯一的统计均衡。

这个定理意味着，如果满足这四个假设，那么改变初始状态、历史和干预措施，都不能改变长期中的均衡。如果各个国家根据某个固定的概率在独裁统治和民主制度之间变换，那么对其中的某些国家强加民主制度，或对它们进行干预、鼓励民主化，从长期的角度来看不会产生什么影响。如果主流的政治意识形态的变动满足上述四个假设，那么历史也不能影响意识形态的长期分布。如果一个人的精神状态可以用马尔可夫模型来表征，那么鼓励和支持的话语都不会产生长期影响。

佩龙-弗罗宾尼斯定理（**Perron-Frobenius Theorem**）

一个马尔可夫模型必定收敛于一个唯一的统计均衡，只要它满

足如下四个条件：

状态集有限：$S=\{1，2，\ldots，K\}$。

固定转换规则：状态之间的转移概率是固定的，即在每个周期中，从状态A转换为状态B的概率总是等于P（A，B）。

遍历性（状态可达性）：系统可以通过一系列转换从任何状态到达任何其他状态。

非循环性：系统不会通过一系列状态产生确定的循环。

需要强调的是，从佩龙-弗罗宾尼斯定理中得出的结论不应该说明历史是不重要的，而应该是：如果历史确实是重要的，那么必定会违背模型的其中一个假设。有两个假设，即状态集有限和非循环性，几乎总是成立的。遍历性似乎有可能会被违背，就像当盟国发动战争并且不能转变为联盟时，会出现非遍历性。尽管确实有一些这样的例子，但是遍历性通常也能成立。

状态之间的转移概率是固定的这个限制是最有可能被违背的假设。因此，这个模型表明，如果历史确实是重要的，那么必定存在某种潜在的结构因素改变了转移概率（或者改变了状态集）。考虑一下帮助贫困家庭摆脱困境的扶贫政策。事实证明，导致社会不平等的那些因素通常不会受到政策干预的影响。[3]同时，马尔可夫模型又表明，改变家庭状态的政策干预措施，例如旨在帮助成绩落后的学生的特殊帮扶计划，或者食物募捐活动，只能在短期内带来改善，不会改变长期均衡。相比之下，提供资源和培训，以提高人们保住工作的能力，进而减少从就业变为失业概率的干预政策则有可能会改变长期结果。无论如何，马尔可夫模型至少为我们提供了一些术语，使我们能够理解状态与转移概率之间的区别。它也告诉我们一个基本道理——与其改变当前状态，还不如改变结构因素，而后者更有价值。

销量-耐久性悖论

销量-耐久性悖论（sales-durability paradox）说的是，产品或创意的流行程度与其说取决于它们的相对销量，不如说取决于它们的耐久性。只需要将拥有某种类型商品的人的比例设定为状态，就可以用马尔可夫模型来解释这个悖论。在这里我们考虑两种不同的地板，一种是瓷砖（耐用品），另一种是油毡（销量更大的商

品）。当销量更大的商品（在这里这个例子中是油毡）不那么占主流时，就会产生这种悖论。

在我们的模型中，假设油毡的销量是瓷砖的3倍。为了刻画耐久性差异，假设每年有1/10的人必须更换他们的油毡地板，而需要更换瓷砖的人则只有1/60。在由此而得到的马尔可夫模型的均衡中，有2/3的地板都是用瓷砖铺就的。[4]

销量-耐久性悖论背后的逻辑，也可以用来解释市场份额与品牌忠诚度（某人改用其他品牌的产品的可能性）之间的正相关关系。在马尔可夫模型中，更低的品牌忠诚度在均衡状态时必然意味着更低的市场份额，因为忠诚度所起的作用就像耐久性一样。这种经验规律有时也被称为“祸不单行法则”（double jeopardy law）。如果一个企业的产品的品牌忠诚度较低，那么其销售量往往也较低。[5]

一对多的马尔可夫模型

马尔可夫模型可以应用于各种各样的环境。我们可以用马尔可夫模型来对四种核酸之间的遗传漂变进行建模分析：腺嘌呤（A）、胞嘧啶（C）、胸腺嘧啶（T）和鸟嘌呤（G）。如果每种核酸都有很小且相等的概率成为其他三种类型之一的核酸，就可以写出一个刻画遗传漂变的转移矩阵。

我们也可以用马尔可夫模型对身体健康演变的轨迹进行建模，方法是将不同的健康类别（如优秀、中等和糟糕等）设定为不同状态。这样的模型可以用来评估药物治疗、行为改变和手术等健康干预措施如何转移概率和均衡分布。那些能够产生更好均衡的健康干预措施是值得追求的。[6]

马尔可夫模型还可以用于识别国际危机的不同模式，并能够用于区分会导致战争的过渡与会带来和平的过渡。[7]不过，在这个领域的应用要求我们估计两种不同的模型：一种模型中，危机导致了战争，另一种模型中，战争爆发前实现了和解。如果这两个模型中的转移概率有显著差异，那就可以对现有的各种模式进行比较，例如轰炸、劫持人质、不交换囚犯以及逐步升级的强硬姿态等，然后看哪个过程对数据的拟合更优。

这种通过马尔可夫模型将不同模式区分出来的方法，还可以用来辨

别书籍或文章的作者。只要事先已经确定了某个作者的某些已知著作，就可以估计出一个词之后出现另一个词的概率。例如，在本书中，单词the跟在for后面的次数，相当于单词example跟在for后面次数的四倍。我们可以将这类信息表示为一个大转移概率矩阵。本书的矩阵看上去肯定与其他人的书的矩阵有所不同。如果为Melville（梅尔维尔）和Morrison（莫里森）分别构建单词转移矩阵，那么我们肯定会看到它们在单词对之间的转换是不同的。[8]

运用这种方法，我们可以通过构建模型来确定《联邦党人文集》（*Federalist Papers*）所收录各篇文章的作者。《联邦党人文集》共收录了85篇文章，分别由亚历山大·汉密尔顿（Alexander Hamilton）、约翰·杰伊（John Jay）和詹姆斯·麦迪逊（James Madison）在1787年和1788年写成，目的是说服当时的纽约州民众支持美国宪法。每篇文章的署名，都用了统一的笔名“Publius”。尽管大多数文章的真正作者都已经确定了，但仍有少数几篇文章一直存在争议。

一项利用马尔可夫模型来辨别作者的研究认为，所有有争议的文章都是詹姆斯·麦迪逊写的。[9]即使汉密尔顿或杰伊写了这些文章，他们也是在模仿麦迪逊的风格。对哲学家阿琳·萨克森豪斯（Arlene Saxonhouse）发表的4篇论文和未发表的12篇短篇论文的分析结果表明，其中至少有3篇论文都有很高的概率可以归入霍布斯名下。[10]当然，在这两个例子中，模型都不一定能给出准确的答案。但是，这些模型确实生成了知识。我们要依靠自己的智慧做出判断：对这个模型与其他模型或直觉结论，该如何进行权衡。

我们要讨论的最后一个应用，是谷歌公司如何运用马尔可夫模型构造谷歌最初的网页排名（PageRank）算法。网页排名是万维网的搜索方式。[11]万维网由链接连接起来的网站组成，为了估计出每个站点的相对重要性，可以计算一个站点连入和接出的链接数量。在图17-3所示的站点网络中，站点B、C和E各有两个链接，A有一个链接，D没有链接。这种方法提供了对重要性的粗略估计，但是它有很大的缺陷。站点B、C和E都有两个链接，但是站点E看上去似乎比站点B更加重要，这是由它在网络中的位置决定的。

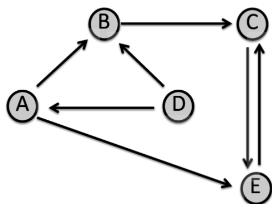


图17-3　万维网上站点之间的链接

网页排名算法将每个站点都视为马尔可夫模型中的一个状态。如果两个站点共享一个链接，那么就在这两个站点之间分配一个正的转移概率。我们暂且为任何链接分配相等的概率，也就是说，假设在A上的搜索者有同样的可能性移动到B或E上。如果搜索者来到了E上，那么将永远交替出现在C和E之间。或者，如果搜索者选择了B，那么他还是会去C，然后再一次开始在C和E之间交替出现。实际上，从任何站点开始都会导致交替出现在C和E之间这个结果。我们发现C和E似乎是更加重要的站点。然而不幸的是，这个模型不满足佩龙-弗罗宾尼斯定理的两个假设。该系统无法从任何站点到达任何其他站点：无法从C到达D，转移概率在C和E之间创建了一个循环。

为了解决这两个问题，谷歌公司的算法加入了一点：从任何站点都能够以一个很小的随机概率移动到任何其他站点，如图17-4所示。现在，这个模型就满足佩龙-弗罗宾尼斯定理的所有假设了，而且存在唯一的均衡。于是，所有站点都可以根据它们在那个均衡中的概率进行排序。一个从A开始的搜索者，最有可能在几次搜索后以到达C或E结束。一旦到达C或E之后，他将会在这两个站点之间反复来回，直到尝试前往一个随机站点为止。如果他到了A或者D，那么回到C的路径很可能会经过B或者E。因此，B的排名应该高于A或D。图17-5所示的唯一统计均衡表明，这个结论是正确的。A、B和D都很少被访问，其中，B的访问量最大。

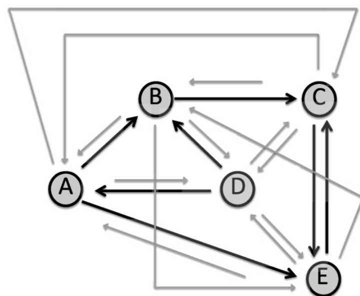


图17-4　在站点之间添加随机移动

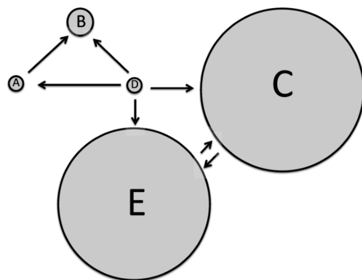


图17-5　网页排名模型的统计均衡

网页排名可以看作随机游走与马尔可夫模型的组合。如果将网页排名视为一种算法，就会发现可以用它来生成任何网络的排名。我们可以让节点代表棒球队或足球队，再用转移概率表示一支球队击败另一支球队的时间百分比。[12]如果球队之间只打一场比赛，那么可以根据胜率来分配转移概率。由此而得出的排名虽然不是最终的，但却是对专家主观评估意见的有益补充。我们还可以利用食物链数据，通过网页排名算法来计算物种之间的相对重要性。[13]

小结

马尔可夫模型描述了以固定的转移概率在不同状态之间转换的动态系统。如果再假设这个过程能够在任何两个状态之间转移，并且这个过程不会产生循环，那么马尔可夫模型就可以得到唯一的统计均衡。在均衡中，人或实体在各个状态之间移动，但是各个状态的概率分布不会发生改变。由此可见，当一个过程接近均衡时，概率的变化就会减弱。用曲线图表示，就表现为曲线的斜率走平。回想一下，我们在讨论线性模型时对美国加利福尼亚州人口增长的讨论。加利福尼亚州的人口增长已经放缓，因为随着人口的增长，离开加利福尼亚的人数也在增加。即便离开的人数所占的比例没有发生变化，这个结论也是成立的。

在应用马尔可夫模型解释现象或预测趋势时，建模者对状态的选择至关重要。状态的选择决定了这些状态之间的转移概率。一个简单的关于药物成瘾行为的马尔可夫模型，可能只需要呈现两种状态：或者是药物成瘾者，或者是正常的人。而更精细的模型则可以根据使用频率来区分药物成瘾者。无论对状态的选择如何，如果上面的四个假设都成立（并且在这种情况下，关键检验将变为转移概率是不是能够保持固定不变），那么系统将会存在一个唯一的统计均衡。系统状态的任何一次性变化都最多只能产生一些暂时性的影响，减少均衡中的药物依赖必须改

变其转变概率。

按照同样的逻辑，我们可以推断，那些试图通过为期只有一两天的活动来激发学生学习兴趣的做法，可能不会产生什么有意义的影响。与此类似，进入社区“送温暖”、来到公园“捡垃圾”的志愿者也可能无法带来什么长期收益。任何一次性的资金涌入，无论其规模大小，影响都会消失，除非它改变了转移概率。2010年，马克·扎克伯格（Mark Zuckerberg）向新泽西州纽瓦克市的公立学校捐赠了1亿美元，并吸引了不少跟风捐赠者。这种一次性捐赠，尽管摊到每个学生头上达到了大约每人6 000美元，但对考试成绩却几乎没有产生任何可衡量的影响。[14]

马尔可夫模型是通过区分以下两类政策来指导行动的：一类政策能够改变转移概率，而改变转移概率可以产生长期影响；另一类政策只能改变状态，并且只能产生短期影响。如果转移概率无法改变，那么我们必须定期重置状态才能改变结果。沉溺于辛劳工作可能会产生导致好强、自私和压抑的心理状态转移概率，而每天锻炼、冥想或参加宗教活动则可能帮助人们以一个感恩的、富有同情心的、放松的心理状态迎接每一天。周末休息也有类似的功能，已婚夫妇不时过一过约会之夜也有很好的效果。这两者的共同作用是，能够暂时使一个人的状态远离均衡。

当然，并不是每个动态系统都满足马尔可夫模型的假设。在不满足马尔可夫模型假设的情况下，历史、干预政策和事件都可能会产生长期影响。例如，在波利亚过程中，结果改变了长期均衡。对系统的重大干预或冲击可能会改变转移概率甚至是整个状态集。蒸汽机、电力、电报或互联网等重大技术变革，改变了经济的可能状态集。重新界定权力架构或制定新政策的政治和社会运动，也会改变状态集。因此，我们也许更应该将历史视为一个马尔可夫模型序列，而不是视为一个向不可避免的均衡方向发展的过程。

马尔可夫决策模型

马尔可夫决策模型（Markov decision model）是对马尔可夫模型的一种修正，方法是将行动包括进来，行动会带来回报，而回报则以状态为条件，还会影响状态之间的转移概率。考虑到行动对转移概率的影响，最优行动并不一定是能够最大化即时回报的那个行动。

例如，要在上网与学习这两个行动之间做出选择。上网总能带来相同的回报。而当学生选择学习时，则有两种可能，既可能觉得充实，也可能觉得无聊。如果觉得充实，学习就可以获得高回报，如果觉得无聊，学习就只能获得低回报。

为了加入行动对转移概率的影响，假设一个觉得学习无聊的学生转为在上网时，仍然会处于无聊状态；而一个觉得学习充实的学生转为在上网时，有一半的时间会变得无聊。假设一个学习的学生有75%的机会在下一个时期处于觉得充实的精神状态，而无论他当前的状态如何。于是：

行动：上网（U），学习（S）

状态：觉得无聊（B），觉得充实（E）

奖励结构

	无聊（B）	充实（E）
上网（U）	6	6
学习（S）	4	8

转移映射

	无聊（B）	充实（E）
上网，无聊（U，B）	1	0
上网，充实（U，E）	1/2	1/2
学习，无聊（S，B）	1/4	3/4
学习，充实（S，E）	1/4	3/4

马尔可夫决策模型的解决方案由每个状态下采取的行动构成。之前讨论过的短视最优反应行为，在每个状态下都选择能够最大化奖励的行为。在现在这个例子中，这种选择对应于无聊时上网、精神充实时学习。

但是，这种短视的解决方案会导致学生陷入无聊状态。一旦发生了这种情况，他们就会选择上网，并在所有剩余时间内一直保持

无聊状态。因此，他们的长期平均回报等于6。而总是选择学习的解决方案则会在他们75%的时间里处于充实状态，只在25%的时间里处于无聊状态，从而得到的长期平均回报为7。这个解决方案产生了更高的平均回报，因为他们更多地处于充实的精神状态。

正如这个例子所表明的，将一个决策问题表达为一个马尔可夫决策模型，可以告诉我们更好的行动是什么。通过考虑行动对状态的影响，我们会做出更明智的选择。晚睡与早起和锻炼相比，会产生一个更高的直接回报，购买昂贵的咖啡比自己动手制作咖啡产生更高的回报。然而，从长远来看，我们可能会更乐于坚持锻炼和节省咖啡钱。那么，我们需要一个模型吗？不一定。相反，我们也许只需要时时记起《圣经·箴言》21:17就可以了：“爱宴乐的，必致穷乏；好酒爱膏油的，必不富足。”这可能是对的；但是我们同时可能记得《圣经·传道书》8:15所说的：“我颂赞喜乐，因为世人在天日之下再好不过的，就是吃喝欢乐。”是的，我们总能找到一对相反的谚语。通过将我们的选择嵌入马尔可夫决策模型中，可以使用逻辑来确定在给定的情境下，哪些常识性的建议真的有用。

18　系统动力学模型

我们尚未充分理解支配系统行为的原理。

杰伊·赖特·福雷斯特（Jay Wright Forrester）

在本章中，我们将介绍一系列系统动力学模型。[1]这些模型可以分析那些有反馈和相互依赖性的系统，可以用于对生态和经济、供应链和生产过程建模。系统动力学模型提高了我们通过包含正反馈和负反馈的逻辑链进行思考的能力。系统动力学模型通常要包括源（source）、汇（sink）、存量（stock）、流量、速率和常数等组成部分。源产生对系统的输入，汇吸收输出，存量跟踪变量的水平，流量刻画各存量水平之间的反馈，速率和常数用于流量，流量可以是随时间变化的，也可以是固定不变的。

系统动力学模型可以同时包括正反馈和负反馈。当变量或属性的增加导致同一变量或属性的更大增加时，就会出现正反馈，第6章中讨论的马太效应就是如此，成功会带来更大的成功，畅销会带来更大的畅销。就学术论文和专利而言，引用会导致更多的引用。

负反馈会抑制趋势。我们必须小心，不要从“负”这个词的字面含义推出任何规范性的含义。负反馈往往能够带来理想的属性，防止泡沫和崩盘。当我们吃东西吃到饱时，大脑会收到停止进食的信号；当企业的利润超过了正常的经济回报时，竞争者就会进入，使在位者的利润减少并防止他们剥削消费者；当一个物种过度繁殖时，它的成员会为食物展开竞争，从而使增长率下降。在这些情况下，负反馈都有助于系统层面的鲁棒性。

使用系统动力学模型，我们通常可以确定复杂性的原因。当系统既包括正反馈又包括负反馈时，就会产生复杂性。在生命游戏中，活元胞会导致新元胞变活，但是过度拥挤又会导致元胞死亡。

这些模型可以将流量和存量水平表示为数学函数。我们可以校准这些定量模型，以便解释过去的存量值、预测未来值，同时估计干预措施的影响。然后，我们就可以利用这些模型解释、预测和指导行动。这些模型也可以是定性的，可以用加号或减号标记每个箭头以表明内在逻辑。[2]

本章内容由五部分组成。在第一部分中，为了引入系统动力学的术语体系，我们先构建了一个定性的面包店模型。在第二部分中，我们构建了一个基于洛特卡-沃尔泰拉方程（Lotka-Volterra equation）的捕食者-猎物模型。我们这个版本假设了存在狐狸和野兔两个相互作用的种群，并同时嵌入了负反馈和正反馈。在第三部分中，我们阐述了如何利用系统动力学模型来预测恶性循环。在第四部分中，我们描述了WORLD3模型，它是一个关于全球经济的巨型模型。最后，在本章的结论部分，我们讨论了系统动力学模型为什么经常会产生违反直觉的结果，这些结果证明了人类直觉推理的局限性和模型作为逻辑辅助工具的价值。

系统动力学模型的各个组成部分

任何一个系统动力学模型都由源、汇、存量和流量组成。源产生存量；存量是某个变量的数量或水平；流量描述了存量水平的变化；汇能够捕获来自存量的流量输出；汇和源是不包含在模型中的过程的“占位符”；存量水平会根据源和流量随时间推移而变化。例如，在一个关于游乐园的系统动力学模型中，公园中的人（一个存量）的人数会随着人们的到来（一个源）而增加，而增加的速度可能又取决于其他参数，例如天气状况、广告投入和门票价格等。

系统动力学模型使用如图18-1所示的表征系统。源和汇用云表示，存量用方框表示，流量用箭头表示并附以加号或减号标识，可变流量用倒三角形表示，而不变流量则用被流量箭头对穿的圆圈表示。正箭头表示正反馈（更多会带来更多），负箭头表示从一个变量到另一个变量的负反馈。

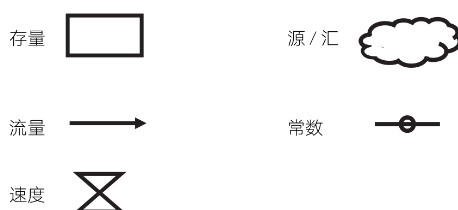


图18-1 系统动力学模型的组成部分

为了让读者尽快熟悉这套符号，在这里先给出一个由面包师、面包和顾客组成的简单的面包店系统动力学模型：面包师制作面包，顾客购

买面包。如果面包师生产面包的速度超过了顾客购买面包的速度，面包的库存量就会增加，面包店将会堆满面包。或者，如果卖出面包的速度超过了面包师生产面包的速度，面包店将永远处于卖光的状态。为了使模型更加真实，我们可以假设让面包师根据面包的存量来调整面包生产的速度，如图18-2所示。在这幅图中，包括了一个从面包库存量到面包师生产面包的速度的流量（一个箭头）。我们在这个箭头上放了一个负号，以表示随着面包库存量的增加，生产面包的速度会下降。如果适当地调整速度，模型将产生一个均衡，使面包生产速度收敛于顾客购买面包的速度。

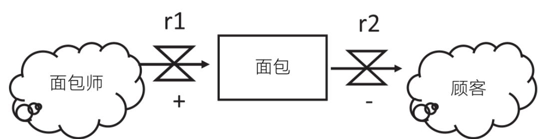


图18-2　一个关于面包店的系统动力学模型

为了使模型更加真实，还可以加入第二个存量排队长度，它等于在面包店外等待的人数。然后再加入第二个源，即潜在顾客，它会增加排队的人数。排队长度较短或适当，可能会吸引更多顾客，而过长的排队长度则会让客户流失，导致他们调头而去。为了刻画排队长度对来自源到达速度的影响的这种可变性，我们在这个箭头上方标记了（+/-）的符号，还在从排除存量线条到顾客购买面包速度的箭头上方加了一个加号，表示排队的人越多，客户购买面包时做决策的速度越快，如图18-3所示。

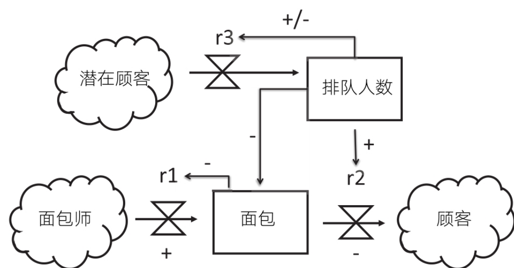


图18-3　一个更加精细的关于面包店的系统动力学模型

这个模型可以根据数据进行校准。我们可以估计出人们根据排队长度加入排队队伍的速度。然后，面包师可以根据面包存量和排队长度确定烘焙的最佳速度。这个速率将提供一个起点，从中可以获得更好的速率。即使没有校准，编写模型的行为本身也增加了价值。面包师很清楚

排队的长度对整体销售量的重要性。

捕食者-猎物模型

我们现在介绍捕食者-猎物模型，这是一个用来刻画野兔数量（猎物）与狐狸数量（捕食者）之间关系的生态模型。这个模型包括两个正反馈：野兔生下野兔、狐狸生下狐狸。它还包括一个负反馈：狐狸吃野兔。该模型假设野兔的存量水平很高，狐狸会产生更多的后代。图18-4定性地表示了这些假设，但没有量化关系。从图中可以看出，随着狐狸数量的增加，野兔数量的减少，从而又导致狐狸数量减少。而随着狐狸数量的下降，野兔数量应该增加，进而导致更多的狐狸。逻辑表明了循环的可能性，也可能是均衡，但我们无法确定。

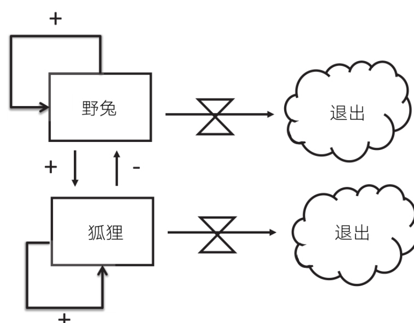


图18-4　捕食者-猎物模型的系统动力与模型

为了更加深入地了解到底发生了什么，我们需要构建出这个模型的定量版本。假设流量是线性的，取决于存量水平。再假设在没有狐狸的情况下，野兔的数量以固定的速度增长，而在没有野兔的情况下，由于缺乏食物，狐狸的数量以固定的速度减少。再假设，在这个模型中，野兔和狐狸相遇的概率与狐狸数量乘以野兔数量的积成正比。为了捕捉发生这种相互作用时吃野兔的狐狸的行为，假设狐狸的数量以某个恒定的速度乘以上述乘积的速度生长，同时野兔的数量则以某个恒定的速度乘以上述乘积的速度减少。由此而得到的方程就是通常所称的洛特卡-沃尔泰拉方程。

洛特卡-沃尔泰拉方程

假设一个生态系统由H只野兔和F只狐狸构成。野兔的数量以g的速度增长，狐狸的数量则以d的速度减少。当野兔和狐狸相遇

时，野兔以a的速度死亡，狐狸以b的速度增长。根据这些假设，可以给出如下微分方程组：[3]

$$\dot{H}=gH-aFH \qquad \dot{F}=bFH-dF$$

从这个微分方程组可以推导出两个均衡，一个是灭绝均衡，即$F=H=0$；一个是内点均衡，即$F=\frac{g}{a}$和$H=\frac{d}{b}$给出。

这两个微分方程分别描述了野兔和狐狸的数量随时间而发生的变化。当方程等于零时，野兔和狐狸的数量不再改变，系统处于均衡状态。第一个均衡，即灭绝均衡（extinction equilibrium），是说野兔和狐狸都不复存在。因此，这个模型预测，在一定条件下，捕食者-猎物这种关系会导致两个物种同时灭绝。当然，不是所有情况下都会发生这种事情，不然的话，地球上也就不会有任何物种存活了。

第二个均衡是内点均衡（interior equilibrium）。在内点均衡中，狐狸和野兔的数量均为正。在这种均衡状态下，狐狸的数量随野兔的增长速度而增加，但是如果狐狸和野兔之间的每次相遇都会导致野兔种群规模以更快的速度减少的话，狐狸的数量就会减少。这两个结果都很直观：如果野兔繁殖更快，那么这个生态系统可以养活更多的狐狸；如果每只狐狸都需要更多的野兔才能活下去，那么这个生态系统就只能养活更少的狐狸。两个结果都符合我们的直觉，我们想要的也正是这样的结果：模型应该能产生符合直觉的结论。

但模型也应该能够产生不那么直观的结论。幸运的是，这个模型也正是如此。它表明，狐狸的均衡数量完全不依赖于狐狸的死亡率。如果狐狸以更快的速度死亡，野兔的均衡数量就会增加，剩下的狐狸的食物就会更加丰富，而这就意味着狐狸的增长速度更快。狐狸的快速增长速度恰好抵消了狐狸更高的死亡率。

类似的逻辑同样适用于野兔种群。野兔的均衡数量不依赖于野兔的增长速度，也不依赖于野兔被狐狸吃掉的速度，相反，野兔的数量取决于狐狸的死亡速度和狐狸将野兔吃掉"转化为"更多狐狸的速度。对于这些结果，直觉"失败"了，因为我们无法直观地想清楚这里面的反馈机制。野兔增长速度提高的直接影响是会有更多的野兔，间接影响是会有更多的狐狸，而这又意味着更少的野兔，这两种效应会相互抵消。能够得出像这样的非直观结论，正是系统动力学模型的标志特征。直觉在这里"失败"了，因为我们只锁定了直接影响而未能思考整个逻辑链。即便

增加（或减少）速度或流量的直接影响是增加（或减少）存量，系统以正反馈和负反馈的形式产生的影响也意味着，其他存量的值也会发生改变，因此速度或流量变化的净效应可能会减少、会被抵消，甚至可能会被逆转。

通过数学演算，我们可以证明洛特卡-沃尔泰拉方程存在这两个均衡，但是并不知道这两个均衡当中哪一个会实现。确实，如果模型在一开始就处于某个均衡状态，那么它将保持在那个均衡状态。但在运行模型之前，我们并不知道这个方程组是否将会产生均衡、周期性、随机性，抑或复杂性。我们所知道的只是，均衡的确存在。

对方程的模拟会产生滞后循环（lagged cycles）。首先，其中一个物种变得“人口众多”，然后，它的数量开始减少，另一个物种的数量开始增加。实证研究表明，这种循环确实很常见。图18-5显示了位于苏必利尔湖中的一个纵长72千米的名为罗亚尔岛（Isle Royale）的岛屿上，狼（捕食者）的数量与驼鹿（猎物）的数量50年来一直处在此消彼长的过程中。不难注意到，捕食者和猎物物种规模的波动，都呈现出了滞后循环的特征。由于模型省略了地理因素、其他物种的影响、天气的变化、物种内部的异质性等，因此其模式并不像我们预期的那么规则。

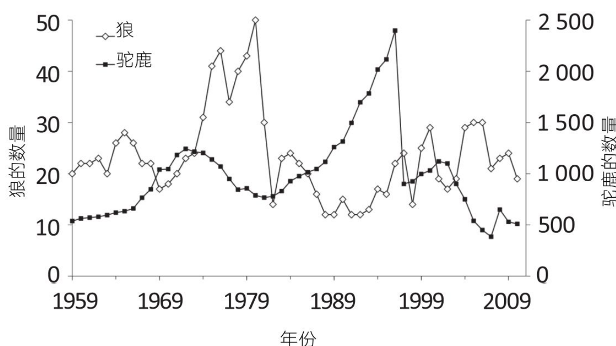


图18-5　罗亚尔岛上的狼和驼鹿

对洛特卡-沃尔泰拉方程的上述分析进一步支持了我们早先的观察结论，也就是不应该将均衡的存在与均衡的实现混为一谈。在图中所示的这种情况下，系统产生的是循环而不是均衡。不过，动态循环是围绕着均衡发生的。因此，均衡可以告诉我们，狐狸和野兔的数量平均来说是多少。由此可见，我们早先得到的违反直觉的结果，也就是提高狐狸（或野兔）的增长速度对狐狸（或野兔）的均衡规模没有影响，在总体层面上仍然是成立的。

系统动力学模型的应用

系统动力学模型既可以包括正反馈回路，也可以包括负反馈回路。正反馈回路会导致良性循环，例如，当两个国家之间的信任度提高，会导致贸易上升、军事对抗减少，从而进一步增进信任。但是，正反馈回路也可能导致恶性循环，例如，某个地区就业机会的减少，可能会削弱人们学习掌握技能的动力，而这反过来又可能导致企业由于缺乏合格的劳动力而迁离这个地区，进而进一步减少工人去努力获得技能的动力。

系统动力学模型可以帮助我们预测恶性循环。2008年，许多国家的经济都面临着严重的金融压力。当资产价格下跌时，过度杠杆化的银行离破产只有一步之遥，投资者和储户开始担心他们投资的安全。有些国家，如美国，只在一定限度内对银行存款进行保险；而在其他一些国家，例如在澳大利亚，则根本不存在存款保险。

为了防止恐慌，澳大利亚决定引入存款保险制度。逻辑上看似乎是合理的：对存款保险可以防止银行挤兑。但是，这种想法其实只考虑到了系统的一部分。在引入存款保险制度的过程中，这种制度本身的一个缺陷会导致悲剧性的后果。如果把相关的系统动力学模型写出来，这一缺陷就会变得很明显。在关于金融系统的系统动力学模型中，每个银行（都是一个存量）都有一定的资产规模。储户将钱存入银行并获得一定回报，银行的借款人用这笔钱进行投资。存款保险制度保证存款人放在银行的资金的安全。

此外，人们还会将资金投入股票市场和货币市场基金。每种类型的投资都是存量。一旦我们开始在各种存量之间绘制箭头，即存量之间的流量，存款保险这种制度的缺陷就会变得很明显。存款保险制度的直接影响能够提高银行的安全性，从而使银行更具吸引力（图18-6中的箭头1）。但是，它同时还会使其他类型的投资变得不那么有吸引力。想象一下，假设在那个动荡的金融危机时期，你自己是一个投资者，在银行和货币市场上都投入了一定资金。现在，你的银行存款已经投保了，但是你的货币市场基金则没有。因此，谨慎的行动是增加银行存款（箭头2）并退出货币市场（箭头3）。

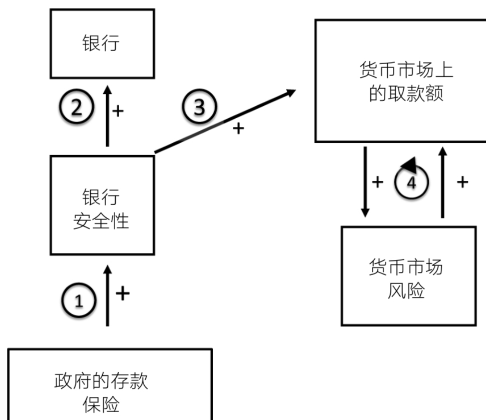


图18-6　金融系统的系统动力学模型

随之而来的是一个恶性循环：货币市场投资的减少使它的风险增加，而风险增加又导致更多的资金从货币市场撤出，从而形成了一个正反馈回路（图18-6中的循环路径4）。撤出资金提高了风险，而风险的上升又会导致更多的提款，从而导致更大的风险。这个政策似乎肯定会造成货币市场的崩溃，现实也确实如此。在决定为银行存款提供保险的短短四天之后，政府就不得不冻结了货币市场上的所有账户，不然整个行业都会彻底崩溃。显然，这个决定带来了灾难性的后果。数以百万计的退休人员，都需要从这些账户中提取资金去购买食品、交纳房租，支付其他基本必需品。[4]

虽然从“事后诸葛亮”的角度来看，这种恶性循环似乎是显而易见的，但是我们其实无法保证，即使澳大利亚的决策者当时已经构建了一个系统动力学模型，他们明白采取这项政策会带来的恶劣后果。重要的是，写出这样一个模型，至少有可能会帮助他们看到，银行存在保险制度会在金融体系内产生更加广泛的影响。如果是这样，他们应该是有可能会注意到这个恶性循环的。这个例子还说明了数据的局限性。来自其他国家的数据表明，保险存款制度有助于稳定金融体系。但是，在这些国家，存款保险制度并不是在金融危机期间创立的，因此这种经验数据对澳大利亚当时的政策选择没有意义。

WORLD3模型

接下来，考虑一个更复杂的涵盖了全球经济的系统动力学模型，这个模型通常被称为WORLD3模型。WORLD3模型最早出现于20世纪70年

代，当时这个模型预测世界经济行将崩溃，除非各国政府改变自己的经济增长政策和环境政策。[5] WORLD3模型包括了在一个共同的框架内以不同速度增长的多个彼此互动的过程，因而利用它，决策者能够看到各种各样的相互依赖关系。[6]但是，主流经济学家往往不看重WORLD3模型，因为在他们看来，这个模型一方面过于复杂，另一方面又没有考虑到经济行为主体的理性反应。

这个模型假设，人口和经济产出每年以某个不变的百分比增长，但是经济产出会造成污染。随着时间的推移，土地的生产力下降，人口规模将超过经济提供商品的能力限度，那时世界经济就会崩溃。这个预测不禁让我们联想到托马斯·马尔萨斯在差不多两个世纪前提出的那个可怕的警告。

这个模型包含了大约150个变量、300个方程和500个参数，生育率、经济增长率和土地使用率等无不包括在内。要校准这个模型，必须根据数据估算出这些参数的增长率。WORLD3模型还包括了变量之间的相互作用，这意味着多个参数的变化通常会产生非线性效应。因此，检验模型的鲁棒性需要不断尝试同时更改两个变量、三个变量的多种组合。500个参数可以产生超过12 000个参数对和超过2 000万个参数三元组。对任何人来说，这无疑都太多了。

这个模型的预测是，到2100年，全世界人口将会减少到40亿。不过，约翰·米勒（John Miller）发现，只要调整两个参数——分配给消费的工业产出比例和女性的生育周期长度，就能够使模型预测的世界人口几乎增加一倍，即达到74亿。正反馈带来了巨大的增长。更长的生育周期意味着更多的孩子，他们又需要更多的食物。增大食物在产出中的份额可以让更多儿童生存，生存下来的女性有更长的生育周期，于是生下来的孩子更多。最终结果将是人口大规模增长。[7]

参数的小小变化，就会导致人口增加一倍，这个结果会带来很多麻烦。不过，结果依赖于参数这个事实本身并不是一个弱点。相反，构建这个模型的目的就是为了指导行动、帮助识别有效的政策。例如，这个模型证明，降低生育率能够降低人口增长速度（事实证明确实如此）。再者，由于这个模型已经经过了校准，因此它还可以给出人口增长速度到底可以减少多少的估计值。因此，我们可以将这个模型包括在一个模型集合中，以便得到更加准确的预测。

随着时间的推移，这个模型最初的预测变得不那么准确了，部分原因是随着人口的增长，人口增长率已经放缓，它们不再符合模型的假

设，这正是经济学家早就预料到的一种适应性反应。[8]虽然WORLD3模型的倡导者接受了这种批评，但是他们也指出，这个模型的许多预测，包括与经济增长和世界总人口有关的预测都是非常准确的。至于生育率的降低，这个模型的倡导者认为，如果WORLD3模型的失效是它自己的原因所导致的——因为这个模型使人们认识到了人口过剩问题的严重性和环境的重要性，他们对这种错误的态度是乐见其成的。

小结

在构建系统动力学模型时，我们选择了关键的组成部分（存量），描述了这些部分之间的关系（流量）；然后我们用系统动力学模型进行模拟实验，以发现它的含义。这些模型与马尔可夫模型的不同之处在于，在这里速度是可以调整适应的（它起到了转移概率的作用）。因此，这种模型不一定会达到均衡。我们必须运行模型才能看清楚会发生什么。此外，因为不必求解结果，所以也不用担心假设的易处理性。

系统动力学模型可以包含许多变量，而且可以包括这些变量之间任何类型的反馈。我们当然也可以构建不存在这些东西的模型，但是一旦绘制出了定义存量的方框，建模者就几乎无法在它们之间绘制出一些箭头来。建模者觉得自己有义务追问："还有哪些其他变量也可能会受到影响？这些变量的变化又如何反馈回当前的模型中？"这些问题会带来更加精细的模型。

当然，这种灵活性也可能是需要付出代价的：创建的存量和流量越多，模型就会变得越难以理解。构建一个有用的系统动力学模型的"艺术"体现在，既要包括足够多的细节来揭示我们的直觉哪里行不通，但是也不能包括过多的细节，以至于会创造一个像现实世界一样混乱的泥淖。最有用的系统动力学模型都位于那个边界上。这些模型可以揭示出意想不到的效应，并有助于更好地采取政策行动。正如在本章前面的内容中看到的，即使是意图极好的政策，例如澳大利亚政府推出的存款保险制度，也可能会产生不良后果。

系统动力学模型还阐明了负反馈回路会怎样限制干预政策的效果。强制要求汽车增加安全功能的法律，例如安装防抱死制动系统或安全气囊，可能会导致人们更加鲁莽地开车，进而导致更多人受到伤害。道路加宽可能会导致更多人搬到郊区，从而进一步加剧拥堵。减少香烟中的尼古丁含量则可能会导致吸烟者消费更多的香烟。为性传播疾病开发更

好的治疗方法可能会使人们更容易做出不安全的性行为。这样的例子不胜枚举。[9]从事后回顾的角度来看，这些负反馈中有许多似乎都是显而易见的，但是要想提前做出预测，可能就不是一件容易的事情了。写出一个定性的系统动力学模型能够使这些反馈回路变得清晰，从而帮助我们成为更好的思考者。

系统动力学模型会鼓励我们把反馈包含进模型中，这个事实恰恰是这种方法的优势所在。1696年，英国国王威廉三世推出了一项宅基地税，规定每栋房子的基本税率为两先令，另外还根据窗户的数量收取额外的税收：窗户数量在10扇至20扇的房子要额外缴纳4先令的税，窗户数量超过了20扇的房子则要额外缴纳8先令的税。国王之所以要对窗户征税，是因为它们是可观察的（可以客观地加以测量），并与住房价值相关。如果国王依赖对房子不动产价值的评估，就会出现很多偏袒和贿赂现象。这种对窗口征税的方法是一个好主意。因此在接下来的一个世纪里，这个方法传播到了法国、西班牙和苏格兰。直到1926年，法国才最终废除了窗口税。

作为模型思考者，我们有理由预期有目的会适应的人们会对税收政策做出反应。当然，他们确实以各种各样的方法做出了反应。有些人把房屋上原有的窗户用砖头砌上了。聪明的建筑师很快就提出新的住房结构设计。在这种税收期间建造的许多中产阶级住宅，二楼卧室都没有窗户。在爱丁堡，有一大排房子的卧室完全没有窗户。[10]于是税收收入随之下降了。在这里，坎贝尔定律再一次发挥了作用：政客们制定了一个制度，但是人们找到了“绕过它”的方法。更精细的系统动力学模型应该还会考虑到窗口减少带来的影响。它将包括从被称为“窗口”这种存量到其他性质的箭头，例如民众的健康——民众的健康会因缺乏新鲜空气和日照而受到损害。

系统动力学模型的巨大价值部分在于，它们能够帮助我们深入思考自己行动的影响。我们通常都能够考虑到政策的直接影响：对窗口征税能够增加收入；安装防抱死制动器能够挽救生命。但是我们并不一定随时都会考虑到间接影响，也就是各种正反馈和负反馈。这些模型恰恰能帮助我们更清晰、更深入地进行思考。

19　基于阈值的模型

种族、民族以及其他阶层的融合标志着显著的社会不平等，这是民主社会的至关重要的理想。

伊丽莎白·安德森（Elizabeth Anderson）

在本章中，我们将介绍基于阈值的模型。当外部变量超过或低于特定的阈值时，人们的行为所发生的变化，就是基于阈值的行为（变化）。例如，当一个人决定在价格低于100美元时购买一件外套，或者当某项社交活动的参与者人数达到1　000人就加入进去时，就会出现基于阈值的行为。基于阈值的行为很直观，也不难分析，但往往会产生违反直觉的结果，例如当宽容行为反而导致隔离时。基于阈值的模型也可以产生临界点。例如，当一个人加入社交活动的决定取决于已经参与了该项活动的人数时，随着越来越多的人参与该活动，参与者的总人数也超过了其他人的阈值，从而导致了更多的人加入。

本章中讨论的模型也可以归类为基于主体的模型。在计算机领域，基于主体的模型指可以单独为每个主体建模的计算机程序。基于主体的模型比系统动力学模型具有更细的粒度。系统动力学模型用单个存量变量表示整个群体，而基于主体的模型则做到了在空间中定位主体，并且可以包括行为多样性。由此增加的自由度当然可以带来一些优势，但是我们必须记住，过多的细节会与要建模的原因相冲突。例如，在构建一个用来讨论人们如何选择参加某项社交活动的模型时，就不应该对人们大脑中的每个神经元进行建模。最优粒度水平取决于模型的目的。

本章从格兰诺维特（Granovetter）的骚乱模型（riot　model）开始讨论，接着引入这个模型的双重骚乱版本，用来对初创企业的成长过程进行建模。然后介绍两个隔离模型：第一个讨论参加聚会的一群人如何在不同房间之间移动，第二个讨论城市规模上的隔离。接下来，讨论乒乓球模型（ping　pong　model），它包括了负反馈，可以产生循环或均衡。在本章末尾部分的讨论中，我们将重新回到生命游戏，分析双重阈值规则如何导致了正反馈和负反馈的结合。正反馈会产生相关行为，而那正是负反馈所要抑制的。本章还会进一步讨论模型粒度问题。

格兰诺维特的骚乱模型

在基于阈值的模型中，个体根据某个总量变量是否超过阈值而决定采取两种行动中的哪一种。如果变量的值超过阈值，个体就采取一个行动，否则，就会采取另一个行动。我们要讨论的第一个阈值模型刻画的是骚乱和社会活动。在这个模型中，每个人都可以选择参加骚乱或者不参加，如何选择取决于参与骚乱的人数。模型没有预设任何规范立场，这里所说的骚乱或社会运动，既可以指对暴君的正义反抗，也可以指足球流氓毁坏财物的非法行为，模型对这两类情况都适用。

骚乱模型为每个人分配一个阈值。当参加骚乱的人数超过那个阈值时，这个人就会参加骚乱。[1]一开始，只有那些阈值为零的人才会参加骚乱。考虑到本书的研究目的，这里所说的“骚乱”是一般的社会活动而不是暴乱，因此在这种情况下，参加骚乱也无非是指参与聚会。假设第一天，有200个阈值为零的人发动了一场社会活动。第二天，这200人继续参加，于是参与阈值低于200的人也加入了他们的行列。假设第二天新加入的人有500人，那么第三天，阈值低于700的人也会加入。而这可能会涉及好几千人。

骚乱模型

假设有N个人（用i来索引），每一个人都有一个“骚乱阈值”（riot threshold），不妨记为$T(i)\in\{0,1,\ldots,N\}$。在一开始，骚乱总产值为零，即$T(i)=0$的那些人都会参加骚乱。令$R(t)$等于在时间t参加骚乱的人的数量。在时间t上，如果$T(i)<R(t-1)$，那么第i个人就会参加骚乱。

骚乱模型的三种情况

对骚乱模型的分析表明，阈值的多样性至少与平均阈值一样重要。原因可以通过比较如下三种情况来说明（它们都涉及1 000人）。在第一种情况下，每个人的阈值都为10，因此不会发生社会运动。在第二种情况下，有5个人的阈值为零，10个人的阈值为1，其他人的阈值均为20。那么在这种情况下，一开始会有5个人发起社会活动，第二天会有10个人加入，但此后就再没有其他人加入了。

在第三种情况下，这1 000人中每个人都有一个独特的阈值，范围为从0到999。为了方便起见，我们可以根据这些人的阈值将他们编号为0至999号，第i个人的阈值为i。第一天，第0号人发起一场社会活动，第二天第1号加入，第三天第2号加入……如此类推，每天都有一个“新”人

加入，直到所有1 000人都参加这场社会活动为止。

不难看出，第一种情况的平均阈值最低，但是不会发生社会活动，因为没有人的阈值为零。在第二种情况下，虽然有一些人的阈值为零，但是仍然不足以发动一场广泛参与的社会活动。只有在最后一种情况下，社会活动才会成功。

这个模型表明，阈值的总体分布是非常重要的，而不仅仅只有阈值的均值才重要。因此它也说明，要预测哪些社会活动将会成功有很大的困难。这个模型还可以指导行动，它可以告诉那些希望发动社会活动的人们：要取得成功，除了组织一个核心革命团体之外，还需要“发动群众”，也就是“创建”一个愿意加入他们的群体。

骚乱模型有各种各样的变体，可以用来分析人们什么时候会起立鼓掌、人们的观点会怎样变化（例如，会不会接受同性婚姻）、时尚潮流的改变（例如，愿不愿意佩戴蝶形领结），以及市场动态演变（例如，是否投身于股市泡沫或房地产泡沫）。在所有这些情况下，人们的行为都可以通过基于阈值的规则来近似，而且该阈值因人而异。在所有这些情况下，发生大事件的可能性——无论是大规模社会活动，还是厚边框眼镜的大范围流行，都可能在很大程度上取决于阈值的分布而不是阈值的均值。

市场创造和双重骚乱

骚乱模型还可以扩展到用来分析互联网初创企业。这些初创企业能够创造出新的买家和卖家市场。要想创造一个新的市场，初创企业必须创造一群买家和一群卖家。例如，致力于将狗主人与狗保姆匹配起来的网站，就需要让狗保姆和狗主人到网站上注册，对于提供包裹配送、运输或清洁服务的网站来说也是这样。这些网站要想取得成功，必须先“创建”出这样两个群体，而且必须保证这两个群体以大致相同的速度增长，否则，卖家或买家将无法找到匹配的另一方。换句话说，初创企业必须同时制造出两场“骚乱”。

成功的初创公司爱彼迎（Airbnb）提供了一个关于双重骚乱的例子。爱彼迎很好地将试图出租房子的人与寻找短期住处的人匹配了起来。为此，爱彼迎需要创建两个人群：求租房子的人和出租房子的人。只有当爱彼迎的网站有足够多的可供出租的房子时，想租房子的人才会访问他们的网站。因此，需要让愿意出租房子的人到爱彼迎的网站上注

册，并将他们的房子挂牌出租。爱彼迎的前两次发布都失败了，因为将房子在网站上挂牌出租需要房东付出不少时间，需要上传不少图片并填写其他一些信息。在爱彼迎拥有大量租房者之前，没有房东愿意将自己的房源放到他们的网站上挂牌出租。

因此，一方面，爱彼迎需要有足够多的挂牌房源在租房者当中造成一场“骚乱”，也就是让足够多的租房者来到他们的网站求租房子；另一方面，他们也需要足够的租房者，以便在那些想要出租房子的人当中制造一场“骚乱”。爱彼迎能否成功，取决于这两个人群的阈值。相比之下，难度更大的是怎样让房东将房源放到他们的网站上挂牌出租，因为这需要房东付出更多的努力。爱彼迎克服这个难题的方法是：挨家挨户地上门，帮助房东将他们拥有的房子在爱彼迎的网站挂牌出租。一旦完成了这个任务，租房“骚乱”就开始了，随后又出现了挂牌“骚乱”。[2]爱彼迎之所以能够成功地创造出这个市场，是因为它的创始人能够培育出足够数量的初始租户，并以此触发双重骚乱。他们先创造出了一条狗尾巴，然后用狗尾巴拉出了一只狗。

谢林派对模型与谢林隔离模型

接下来要讨论的两个模型都是经济学家托马斯·谢林提出的，谢林用它们来分析种族隔离问题。人们在很多个尺度上按种族和民族隔离，先按不同的国家隔离，在国家内部又按地区隔离。在美国，城市内部还会出现社区层面的种族隔离。对于这些，有人说这是种族不宽容的证据。但是，这个结论与越来越多的跨种族、跨民族通婚的证据相矛盾。同一群人，既然决定不与不同种族的人住在同一个社会，甚至不愿意与不同种族的人在一张桌子上吃午餐，但是他们却选择与不同种族的人结婚，怎么会这样呢？

如果跨种族婚姻中的那些人，和不与不同种族的人共坐一张桌子上的那些人，分别属于不同的社会阶层，那么我们或许可以解释这些事实，但是事实并非如此。跨种族婚姻在所有收入水平上都有发生，同时在美国最顶尖的大学和学院里也可以发现“种族隔离”的午餐桌。幸运的是，谢林的模型可以同时解释这两组事实，它们说明了宽容的行为也会导致种族隔离。

谢林派对模型

第一个是谢林派对模型（Schelling's party model）。在一定意义上，可以认为这个模型是随机游走和骚乱模型的“混搭”。谢林派对模型描述了一个在一幢有两个房间的房子里举行的派对。这个派对邀请的客人可以分为两个类型，这里的类型可以是男性和女性、黑人和白人、西班牙人和澳大利亚人、哥特人和印第安人……但这都不会影响结果。关键的假设是，每个人都能够区分其他人所属的类型。

谢林派对模型

> N个人参加一个派对，每个人都有一个可观察的类型A或B。每个人随机选择两个房间中的一个。在每个时刻，每个人都有p的概率走到另一个房间去。第i个人的宽容阈值为T_i。对于这个人，如果他所属的类型的人在当前房间内所有人当中所占的比例低于这个阈值，他就离开这个房间。

为了理解隔离是如何在这种存在“宽容”的情况下产生的，假设前来参加派对的是20名澳大利亚人（A）和20名巴西人（B）。每个人都是宽容的，只要同一个房间里有25%的人与自己属于相同的种族，他们就会留在这个房间里。再假设一开始，有一个房间里待了12名澳大利亚人和9名巴西人，另一个房间则有8名澳大利亚人和11名巴西人。这样没有人觉得自己必须走到另一个房间去，但是两个房间之间会有随机移动，这种随机移动会改变每个房间内的“种族结构”。在图19-1中，一个房间里待了11名澳大利亚人和4名巴西人。这个配置处于一个临界点上：只要任何一名巴西人离开，巴西人所占的“人口比例”就会降低到25%以下，从而导致其余3名巴西人也选择离开。如果发生了这种情况，就永远也不会有巴西人再进入那个房间。

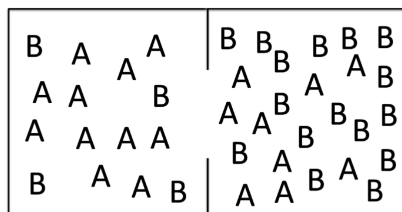


图19-1　一个很容易导致隔离的配置

回忆一下随机游走，我们说过随机游走能够越过任何阈值。一个房间里的澳大利亚人的数量是一维随机游走。因此，如果派对持续时间足够长，那么隔离就是不可避免的。当然，即便是最好的聚会也不一定能

持续得“足够久”，所以并不是所有派对都会以隔离而告终。我们也知道，派对规模越大，隔离结果越不可能出现，因为当参加派对的人更多时，随机游走就必须超过一个更高的阈值。例如，在一个有1 000人参加的派对中，不太可能出现某个房间里某个特定类型的人所占比例低于25%的情况。因此，我们预测，在小型派对中出现隔离的可能性更大。

而且，当参加派对的人有不同的宽容阈值时，我们也应该会观察到更多的隔离现象。要理解个中缘由，不妨假设参加派对的是10名巴西人和10名澳大利亚人，并为每个人分配5%至45%的宽容阈值，使每种类型的人的平均宽容阈值为25%，如图19-2所示。

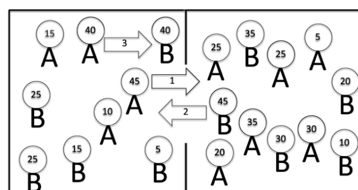


图19-2　异质性阈值导致的“重新安置”

在图19-2中，左边的房间里有5名巴西人和4名澳大利亚人，因此澳大利亚人的比例低于45%，这导致最不宽容的那名澳大利亚人离开这个房间，进入右边的房间（用箭头1表示）。这名澳大利亚人进入了右边那个房间后，就降低了巴西人在那个房间里的比例，导致最不宽容的巴西人换房间（用箭头2表示）。这两次移动使左边房间里澳大利亚人所占的百分比降低到了低于40%的水平，从而导致第二名最不宽容的澳大利亚人也离开左边的房间，进入右边的房间（用箭头3表示）。就这样，一个级联式反应出现了。然而，如图19-3所示，结果不一定是完全隔离，因为最宽容的那些人在任何一个房间里都会觉得舒服自在。在这个模型中，多样化阈值产生了两种不同的影响：它们使隔离更容易发生，同时又使完全隔离的可能性降低，因为非常宽容的人在任何一个房间都不会觉得有问题。

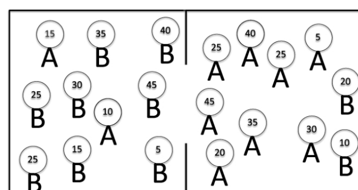


图19-3　异质性阈值导致的“隔离”

这个模型可以用来解释不同职业之间的性别比例差异，例如为什么

会有更多的女性担任护士，更多的男性担任推销员。这种差异也许是由于偏好不同所致，但是只要对这个谢林派对模型稍稍加以扩展，就会发现，如果人们更愿意与相同性别的人一起工作，那么这种不同职业之间的性别比例差异就可能会出现。对于这一点，我们还可以用旋转门模型（revolving-door model）来给出更加正式的解释。旋转门模型有两个基于经验基础的假设：退出某个职业的女性会选择女性更多的新职业；女性退出某个职业的速度略高于男性。[3]例如在生命科学领域中，女性退出生物医学研究的速度比男性更快一些，并且会在雇用了更多女性的医疗保健等职业中谋职，这种行动就会提高这两个职业的性别隔离程度。

谢林隔离模型

谢林隔离模型（Shelling's Segregation Model）将行为主体放在地理空间的不同位置上，就像是放在棋盘的不同格子上（图19-4）。其他方面则都与谢林派对模型相同。这个模型同样假设有两种类型的人，对他们的行为假设也与谢林派对模型相同。

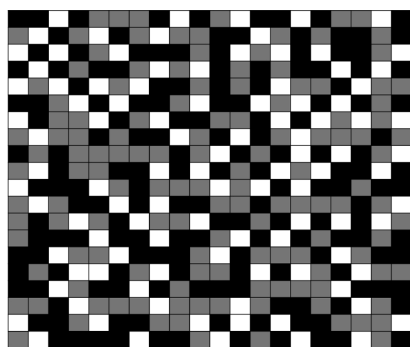


图19-4　谢林隔离模型的初始配置

谢林隔离模型

有N个人，每个人都属于类型A或类型B，随机排列在一个$M \times M$的棋盘上。棋盘上还有部分空间未被占用。第i个人的宽容阈值为T_i。对于这个人，如果他所在方格的8个相邻方格中同一类型的人所占的百分比低于他的阈值，就会重新定位到一个随机的新方格上。

如果个体的平均阈值接近50%，这个模型会产生地域性的隔离，如图19-5所示。出现这种隔离是因为个体只考虑自己所在的“社区”，这种

社区最多只能有8个居住者。几乎任何随机的初始配置都包括了被相反类型的其他人包围的一些人。如果个体进入了具有更多相同类型的个体区域，就可能引起另一种类型的人的重新安置。随着重新安置的不断积累，最终发生了隔离。毫无疑问，阈值多样性会加剧这些影响，原因如前所述，这里不再重复。

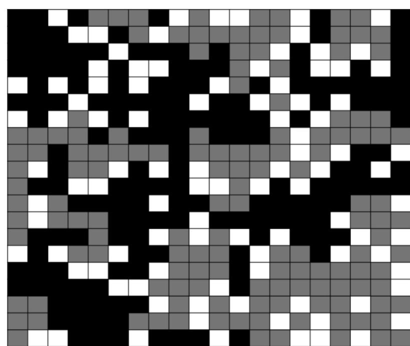


图19-5 谢林隔离模型中“重新安置”后的配置

即使是宽容的人，也会产生隔离的居住模式，这就是托马斯·谢林的开创性名著《微观动机与宏观行为》一书最根本的洞见。

乒乓球模型

现在讨论乒乓球模型，它假设基于阈值的行为会产生负反馈。负反馈可以起到稳定系统的作用，还可以产生循环。乒乓球模型假设了有限数量的实体，它们可以是人，也可以是机械的、生物的或化学的装置。在每个时期，每个实体采取一个正（+1）的行动或一个负（-1）的行动。在第一个时期，每个主体随机选择一种行动，初始状态等于行动的总和。系统所有的后续状态，都等于前一个时期的状态加上所有动作的平均值和一个随机项。每个实体都有一个阈值（从特定的分布中抽取），如果状态绝对值超过了阈值，就选择能够减少状态绝对值的操作。简而言之，如果状态超过了实体的阈值（无论是负的，还是正的），这个实体就尽其所能减少状态的大小。

乒乓球模型

规模为N的种群中的每一个实体随机采取一个初始行动，或者为正（+1），或者为负（-1）。系统的初始状态S_0设置为零。系统的所有未来状态S_t则等于平均行动再加上一个随机变量，即：

$$S_t=\frac{\sum_{i=1}^{N}A_i(t)}{N}+S_{t-1}+\varepsilon_t$$

每个实体i都有一个响应阈值（response threshold）$T_i>0$，它从［0，RANGE］中均匀抽取出来。如果状态的绝对值大小｜S_t｜小于它的阈值，该实体就采取与以前相同的行动，否则，就采取减少状态绝对值大小的行动。即：

如果｜S_t｜$\leq T_i$，那么$A_i(t+1)=A_i(t)$，否则$A_i(t+1)=-\text{sign}S_i(t)$

其中，ε_t是从{-1，+1}中随机抽取出来的。

乒乓球模型有多种应用。人们可能会将时间和资源分配给多个慈善项目。如果某个慈善项目得到了太多的关注或金钱，人们就会开始向其他项目捐钱以保证平衡。例如，一个国家需要维护两个联合国教科文组织世界遗产，人们会在两个遗产上都投入一些时间和资源，但是如果某个遗产受到的破坏更严重，人们就可能会重新分配资源。

正如“乒乓”这个名称所预示的那样，乒乓球模型可以生成围绕着均衡的循环行为。在一个时期内，太多的人选择了这个行为，而在接下来的那个时期，却有太多的人选择了另一个行为。当所有实体的阈值都为零时，在这一个时期内所有实体都会选择这个行为（+1），而在下一个时期内所有实体都会选择另一个行为（-1）。

为了探索阈值多样性如何影响人们的乒乓球式行为模式，还是会找到一个均衡，我们考虑如下三种情况（每一种情况都有100个人参与）。在第一种情况下，假设阈值均匀分布在0到10之间。如果第一个时期中的状态等于6，那么它将超过大约60个人的阈值。这60人中大约有30人将采取行动1，并将改变行动。这些行动总和现在将超过50，因此系统的新状态（前两个时期的平均值）将超过20。这个值超过了所有阈值，因而产生了图19-6所示的乒乓球效应。

在第二种情况下，如果我们增大阈值的多样性，使它们在0到100之间均匀分布，那么乒乓球效应几乎消失。如果再次假设状态在第一个时期等于6，那么平均来说只有6个人的阈值将得到满足。如果这6个人中有3个改变行动，那么状态将趋向于零。这种对偏差的抑制可以从图19-6中看得很清楚（它对应于0到100之间的阈值）。

在第三种情况下，考虑一个适中的范围，让阈值均匀分布在0到60

之间，我们将会观察到更温和的周期，如图19-6所示。因此，在具有负反馈的系统中，阈值多样性会导致稳定性，但是在具有正反馈的模型中则具有恰好相反的效果。

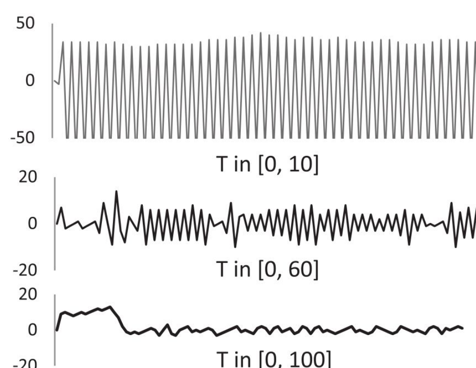


图19-6　三种阈值范围下乒乓球模型的时间序列

模型粒度

反馈的基本逻辑很简单：正反馈强化行动，负反馈抑制行动。只有正反馈的系统会爆炸或崩溃，只有负反馈的系统将稳定或循环，同时具有正反馈和负反馈的系统则可能会产生复杂性。

在系统动力学模型中，反馈作用于存量变量（例如，野兔的数量）和速度（例如，增加面包店的购买率）。基于主体的模型（例如本章中介绍的阈值模型）会产生作为个体行动的结果反馈。这类更精细的模型可能包含多个阈值，阈值多样性会加剧正反馈的影响并缓和负反馈的影响。分布尾部的多样性使骚扰更加容易发生，宽容阈值的多样性则促进了隔离。具有负反馈系统中响应的多样性阻止了同质性反馈可能导致的大幅波动。在企业之间的经济竞争模型中，生产成本的异质性也可以起到类似的作用。随着价格的上涨或下跌，由于各自成本的差异，企业做出的反应会各不相同。[4]

这两种模型之间的差异引发了模型粒度的问题。是创建一个以“野兔”为名的变量（或方框）并描述野兔的数量如何随着其他变量而增加或减少？还是对每一只野兔单独建模？细分变量可以提高描述的准确性，但是模型不是以描述准确性为标准来评分的。请一定要记住统计学大师乔治·博克斯的格言：所有模型都是错的。至于博尔赫斯的地图，虽然与现实世界一样大，却全无用处。许多建模者，包括爱因斯坦，都认为，

我们应该寻求适当的粒度水平。例如，如果要构建一个模型来解释人类的手臂能够施加的力，就不需要考虑DNA。

而在研究社会系统的时候，则可能并不存在理想粒度水平。我们可能需要在多个粒度水平上分别进行探索。通过构建多个模型，每个模型都有不同的粒度水平，然后进行跨模型对话。假设我们试图了解瑞典与芬兰之间的贸易模式，可以这样入手：把两个国家的贸易总量作为变量，并确定宽泛的宏观模式。然后，将每个国家的进出口总额按行业分解，并进一步分解为各个行业内的企业。行业级数据将使我们能够更好地解释过去的模式，并对未来的模式做出更加精确的预测。对各个行业内的企业做进一步深入了解，包括成本结构和增长轨迹等信息，可以得到更好的结果，但是我们需要大量的信息才有可能构建出一个包含这么不断变动着的组成部分的、真正有用的模型。我们甚至还可以对这些企业的领导力进行建模。最后一个层次的拆解式研究，很可能不会带来多少有用的结果，但可能会使一些致力于扩张海外的企业领袖名扬四海。

总而言之，更精细的粒度不一定更好，有的模型可能过于注重细节。我们能够理解更细粒度的模型，也可以从相对粗糙的模型中获益。通过比较不同模型的预测、解释之间的差异，我们可以看清楚假设如何影响结果，也可以看清楚假设的条件性。许多模型不仅在各自包含的变量方面存在差异，而且在何种粒度上分解变量也理所当然地有所不同。

算法骚乱

骚乱模型和乒乓球模型都可以用来解释股市崩盘及随后的价格反弹。我们在此考虑两个案例，它们很能说明一些问题。

第一个案例发生在1987年10月19日星期一，那一天，道琼斯工业平均指数下跌了超过22%。第二天，这场股价崩溃在全球范围引发了连锁反应。这次崩溃的原因直到今天仍然是许多研究的主题。当时，美国经济扩张已经进入了第四个年头。那一年的前8个月，道琼斯工业平均指数的涨幅超过了40%。也许就是因为股票市值的这种上涨，使许多人认为股票价格被高估了。在崩溃前的那个星期天，美国时任财政部部长詹姆斯·贝克（James Baker）威胁说，如果德国再不降低关税，美国就要采取削弱美元政策。这种说法在当天似乎没有引起什么反响。但是第二天，股票市场就崩盘了。14个月后，股票市场又恢复到了崩盘之前的价位。

为了应用本章给出的这些模型，我们可以用一种金融资产来代表整个市场。假设，持有该资产的每个人都有一个“崩溃阈值”（crash threshold）。如果资产价格在某一天的下跌幅度超过了崩溃阈值，投资者就会出售资产，将资金从市场中撤走。这个规则可以捕获趋势交易者或噪声交易者的行为。在此基础上，我们可以构建一个骚乱模型。如果某些投资者在10月19日一早醒来就决定出售大量资产，他们的行为就会导致股票价格下跌。如果股票价格的下跌幅度超出了其他投资者的崩溃阈值，那么其他投资者也会出售，从而导致股票价格呈螺旋式下降。于是在这里就有了一个经典的正反馈循环：抛售导致价格下降，价格下降导致更多抛售。

现在加入乒乓球模型。如果价格下跌得太多，人们会应用第二条规则，换句话说，人们还有一个“便宜货阈值”（bargain threshold），如果价格低于这个阈值，人们就会买入这种金融资产。在这里，投资者是基于价值而非趋势行事。当价格急剧下跌时，便宜货阈值会产生负反馈。买家急于以便宜的价格买入，导致价格停止下跌。

当然，现实世界中的市场要比这个模型（卖家跌破阈值就卖，买家等着机会入货）复杂得多。股票市场包含了许多类型的交易者，包括大型机构、养老基金、各国政府、投机者、投资组合保险公司，以及散户投资者。[5]正因为存在这种多样性，在价格下跌时几乎永远有人愿意买进，从而产生了稳定市场所需的负反馈。

对崩溃的早期分析强调了（计算机）程序交易盛行的影响，这些其实是用计算机程序编码并实施的基于阈值的规则。如果股票市场指数低于某个预先设定的值，就卖出所有股票。诸如此类的规则都是在无人监督的情况下自动执行的。大多数分析师现在认为，程序交易是促成1987年股市崩盘的一个原因，但并不是主要原因。

对1987年股市崩盘更细致的分析表明，大量投资组合保险公司作为必须保证投资者投资组合回报率的交易者，产生了强烈的正反馈，而且这些正反馈并未受到负反馈的抵消。在股票市场大幅下跌的情况下，这些投资组合保险公司会通过抛售股票来止损。随着股票的进一步下跌，这些投资组合保险公司抛售的股票越来越多。实际上，这些投资组合保险公司的行为就好像他们是具有不同阈值的个人组成的群体。有一家投资组合保险公司抛售了超过10亿美元的股票，而10月19日当天全天的成交量只有200亿美元。

第二次案例是一个发生在2010年5月6日下午的闪电崩盘。在三分钟内，道琼斯工业平均指数就下跌了5%。这个闪电崩盘是算法交易的结果。由于现代金融市场的高度复杂性和交易的极高速度，没有人确切知道在闪电崩盘期间究竟发生了什么。但是我们确实知道，有一个庞大的共同基金发出了一个超级大卖单，将超过40亿美元的股票期货倾销到了市场上，而市场上则充斥着各种各样试图利用有利交易的高频交易算法。这种高频算法能够感知价格趋势并开始以极快的速度执行交易，对此，我们可以想象一个有极高速度的骚乱模型。

这就产生了一个“有毒”的市场，交易者担心大型机构投资者知道他们不知道的东西，因此他们退出了市场。[6]鉴于市场行为异常，许多算法停止了交易；但是其他算法则继续销售，然后就发生了闪电崩盘，所有这一切都在几分钟内完成。20分钟后，便宜货阈值规则开始发挥作用，正如乒乓球模型所预测的那样，价格回升到了（接近于）原先的水平上。

20　空间竞争模型与享受竞争模型

> 我们的理论是，如果你需要用户来告诉你，你正在销售的是什么，那么你其实不知道你正在销售的是什么，而且这不可能给客户留下好的体验。
>
> 玛丽莎·梅耶尔（Marissa Mayer）

在本章中，我们将介绍关于个人选择的模型，人们选择时所针对的备选项由其属性来代表。开发这些模型主要是为了刻画消费者的选择行为。买房子的人要考虑它的面积、卧室和浴室的数量以及建筑质量。这些模型也适用于招生的考官、选聘董事长的董事会，以及参加投票的选民，他们都要在申请人或候选人之间做出选择。招生考官要考虑申请人的SAT成绩、平均成绩和课外活动。选民要根据候选人在教育、基础设施、犯罪和税收方面的立场进行评估。除了帮助我们理解个人选择之外，这些模型还可以解释为什么会这样做出选择，例如，为什么我们有这么多种早餐谷物可供选择。

讨论这些模型时，我们会将某些属性描述为空间属性，而将其他属性描述为享受属性。空间属性，例如夹克的颜色或一片面包的厚度，没有最优值。每个人都喜欢特定“数量”的这类属性：一个购买猪排的消费者有自己喜欢的辣度，一个业余滑雪运动员在滑下斜坡时有自己喜欢的下降角度。产品的属性越接近理想水平，消费者对产品的评价就越高，而且理想水平因人而异：一个人可能比另一个人更喜欢辣一些的猪排。

与此不同，在享受属性上，更多（或者在某些情况下更少）总是意味着更好。人们喜欢智能手机的待机时间更长一些，房间的面积更大一些，皮鞋的鞋底更耐磨一些，自己的汽车更省油一些。不过，在现实世界中，大多数选择都是“混合型”的：人们既会考虑空间属性，也会考虑享受属性。

在本章中，我们假设人们总是会选择他们估价最高的备选项（备选方案）。之所以要这样做是出于第4章中已经提到过的一些原因：在对人类行为进行建模的时候，理性行为提供了一个基准，基于理性假设构建的模型不仅容易处理，能够给出符合唯一性要求的预测，并且这种预测在事情重复发生且利害关系很大时与经验相符。

空间（属性）竞争模型（spatial competition model）和享受（属性）竞争模型（hedonic competition model）在经济学和政治学中早就得到了广泛应用，部分原因在于它们便于用数据检验。[1]在本章中，我们可以体会到这种适用性。我们先从空间竞争模型入手分析。接着将这些模型应用于政治领域，并阐明如何使用这种模型来分析现状效应、议程权力和拥有否决权的人的影响力。然后，我们介绍享受竞争模型及其混合模型，以揭示对价格竞争的若干结论。在这个过程中，我们还将说明，如何在模型中应用数据，根据候选人和法官在相关法案和法律案件上的投票历史来推断出他们的立场。我们还介绍了如何推断出清洁空气或更少的通勤时间等没有明确价格属性的隐含价格。[2]

空间竞争模型

空间竞争模型假设备选项可以用一组属性来定义，消费者则可以用一系列理想点来定义。最简单的空间竞争模型考虑的是只有单一属性的产品。在由数学家、统计学者哈罗德·霍特林（Harold Hotelling）提出的最早的空间竞争模型中，这个属性就是地理位置。[3]

霍特林的模型假设一些消费者分散地分布在海滩上（如图20-1中的圆圈所示），海滩上有两个卖冰激凌的小贩（分别用A和B表示）。每个消费者都会从离自己更近的那个小贩手中购买一个冰激凌。图中的分割点位于这两个小贩之间的中点位置，它决定了哪些消费者会从哪个小贩手中购买冰激凌。分割点左侧的那6个消费者会从A小贩那里购买，而分割点右侧的那7个消费者则会从B小贩手中购买。

考虑到消费者更喜欢离自己更近的商品，我们可以对距离给出一种更加抽象的解释。例如，我们可以想象这两个冰激凌小贩位于同一个地点，但是分别出售乳脂含量不同的冰激凌。因此，同样的数字可以表示B小贩出售冰激凌的乳脂比A小贩出售的冰激凌更加浓稠。根据这种解释，消费者的位置就不再是他们在海滩上的物理位置了，而是他们所喜欢的乳脂含量。

我们还可以继续应用一对多的模型思维，用这个模型去分析政治竞争。唐斯模型（Downsian model）将霍特林模型中的地理空间重新定义为从左到右的意识形态连续系统。于是，我们可以对图20-1重新解释如下：A小贩代表自由派的政治候选人，B小贩代表保守派的政治候选人，圆圈则代表选民的意识形态理想点。然后进一步假设，选民更喜欢政治

立场与自己更接近的候选人。

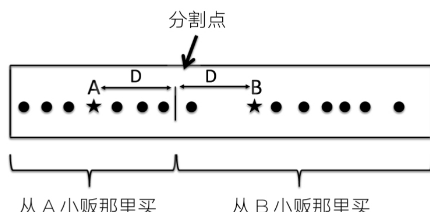


图20-1　海滩卖冰激凌的地理模型

从企业的地理位置和产品属性，转换为政治意识形态，意味着从诸如位置和乳脂含量等物理属性，转换为更抽象的思想观念。虽然我们对于物理属性已经有了明确的测量方法，但是要分配意识形态立场，则需要找到一种新的方法将候选人的行为转化为数字。如果候选人有投票记录，我们只要收集到候选人以往的选票数据，就可以对他们分配意识形态立场。

在这个过程中，我们应该忽视所有缺乏意识形态成分的选票，例如，当所有有权投票的人都一致同意设立全国牛奶日时，这样的投票数据就无须考虑。而对于所有其他投票，则可以依靠专家意见来确定自由派立场和保守派立场。候选人在意识形态区间上的空间位置可以设定为投票给保守派的时间百分比。[4]永远采取保守派立场的候选人将被放置在最右边，一半时间投保守派的票、一半时间投自由派的票的候选人则位于最中间的位置上。

构建了这样一个模型之后，我们就可以通过实证检验来判定美国各政党的意识形态立场的差异是不是变得更大了。如图20-2所示，有一项研究表明，美国民主、共和两党的平均理想点出现了显著且不断加剧的极化趋势。虽然我们不能就此得出美国政党极化程度增加的结论，但是它确实提供了支持这种结论的证据。分析还表明，这种极化主要应归因于共和党的向右倾斜——变得更加保守。

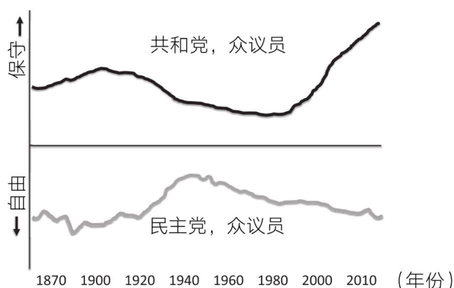


图20-2　美国国会众议院意识形态极化的加剧

增加属性的数量

一般的空间竞争模型可以包括任意数量的属性。比如，沙发可以通过多个物理维度来描述：长度、宽度和深度，结构类型，以及室内装饰类型。消费者从产品中获得的价值（或效用）取决于产品在各个维度上与理想点的距离。我们可以将这种价值函数写为常数项减去备选项与消费者理想点之间的距离。[5]

空间竞争模型

备选项包括N个空间属性：$\vec{a}=(a_1,a_2,...,a_N)$。

做出选择的个体的理想点：$\vec{x}=(x_1,x_2,...,x_N)$。

个体选择了备选项后得到的收益（效用）等于：

$\pi(\vec{x},\vec{a})=C-(x_1-a_1)^2-(x_2-a_2)^2-\cdots-(x_N-a_N)^2$，其中，$C>0$是一个常数。

示例：$\vec{x}=(3,4,6)$，$\vec{a}=(2,1,8)$，$C=20$

$$\pi(\vec{x},\vec{a})=20-(3-2)^2-(4-1)^2-(6-8)^2=6$$

在这种一般的空间竞争模型中，对于两种巧克力棒，可以用可可含量的百分比以及含糖量来表示，如图20-3所示。分割线是连接这两个产品的直线的垂直平分线。理想点位于分割线左侧的消费者更喜欢A，而理想点位于分割线右侧的消费者则更喜欢B。[6]

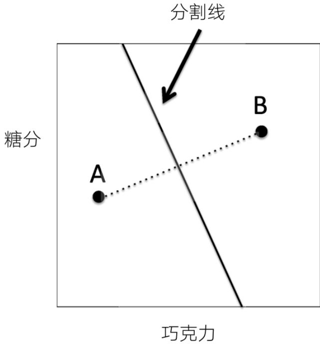


图20-3　两种巧克力棒的产品属性与分割线

这个模型还可以容纳任意数量的产品。要加入第三种巧克力棒，我

们只需在属性空间中增加另一个点即可。为了确定消费者会选择购买哪种巧克力棒，可以再绘制一条分割线，如图20-4所示。这两条分割线可以将理想点的空间划分为三个区域，它们被称为沃罗诺伊邻域（Voronoi neighborhood），领域根据它们与产品的距离划分理想点空间。

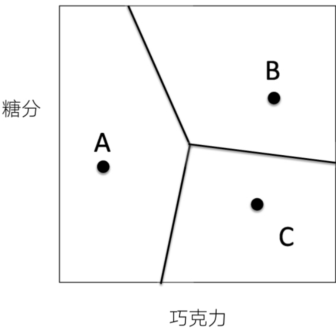


图20-4　有三种产品的空间模型（沃罗诺伊邻域）

唐斯空间竞争模型

接下来，我们应用空间竞争模型来分析政治候选人的意识形态立场“定位”。为了便于讨论，假设候选人都是追求投票最大化，也就是说，他们的首要目标是赢得选举。我们可以从一个简单的例子开始考虑，思考候选人的动机。图20-5显示有13名选民和两名候选人的两种不同情况。回想一下，前面我们讲过，选民更喜欢立场与自己更接近的候选人。在图20-5的上图中，L表示的自由派候选人获得了5票，R表示的保守派候选人获得了8票。在图20-5的下图中，自由派候选人向中间位置移动后，获得了7票，并赢得了选举。

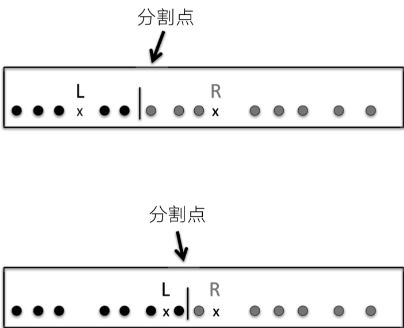


图20-5　向中间位置移动的自由派候选人赢得了选举

自由派候选人有动机向中间位置移动。同样的逻辑，保守派候选人也有这样的动机。也就是说，保守派候选人可以向左侧移动，同时保持在比自由派候选人更靠右的位置上，从而赢得选举。继续按这个逻辑推理下去，自由派候选人也可以选择更接近中间选民理想点的立场。如果继续应用这个逻辑，我们会看到两个候选人应该会趋同到同一点上。这个结果就是通常所称的中位投票者定理（median voter theorem），或称“中间选民定理”。

中间选民定理又可以分别解释强的形势和弱的形势。在强的形势下，中间选民定理意味着各候选人都来到了中间选民的理想点上，采取了相同的立场，但是这显然不符合经验事实。在弱的形势下，中间选民定理意味着候选人有动机采取较温和的立场，经验证据表明这种情况确实会发生。在选举过程中，候选人会趋向中间立场。当然，这种变化不会演变为一场“夺命狂奔”。意识形态信念坚定或受益于位于意识形态光谱极端位置的核心支持者的候选人，对重新定位会非常谨慎。

这个模型将每个候选人和每个选民都归结为意识形态光谱上的一个简单的点，这是一个强有力的假设。捷克作家和政治家瓦茨拉夫·哈维尔（Václav Havel）对这种基于单维意识形态的模型预测非常不以为然，他认为：

> 对于永恒徘徊的人类来说，它（意识形态）提供了一个立即可以安身的家园，所有人必须做的就是马上接受它。突然之间，一切都变得清晰了，生活呈现出了新的意义。所有的奥秘，所有悬而未解的问题，所有的焦虑和孤独感，全都消失了。当然，人们必定要为这个“廉租房”付出沉重的代价。这代价就是放弃自己的理性、良心和责任，因为这种意识形态最重要的一个方面就是向更高权威拱手献出自己的理性和良心。[7]

哈维尔没有说错。我们不应该为了意识形态而放弃理性。模型为我们提供了更好的推理工具，这种特殊的模型有助于我们理解为什么政治家会做出那种行为。利用数据，我们可以确定每个政治家在意识形态的左右区间上的位置的可信度。一个总是采取温和立场的政治家，可以充满信心地站在这个区间中间。

顺便说一句。哈维尔否认人可以简化为一点，即可以用数据进行检验的程度。但是，即便他的批评成立，即便不能根据投票记录来确定一个候选人的意识形态立场，我们也并不一定要放弃这个模型。我们可以

通过赋予他一个区间（而不是一个点）来代表哈维尔所说的意识形态的不确定性。或者，可以为候选人构建一个意识形态时间序列，看他是否保持前后一致。一项对最高法院大法官意识形态立场的研究表明，对有些法官来说，担任法官的时间长了，立场会变得更加自由。[8]

我们还可以增加模型的维度。一个二维模型可以区分社会政策和财政政策。我们可以用这样的模型刻画在社会政策上采取自由主义立场，同时在财政政策上采取保守主义立场的政客。对于美国国会来说，一个维度的模型可以解释大约83%的选票变化，加入第二个维度后，解释力只能增加4%。[9]

除了能够更准确地预测行为之外，增加维度也可能会改变我们的理论结论。我们可以从一个二维模型入手来探讨这个问题。根据对一维模型的分析，我们知道如果第一个候选人在两个议题的立场都不位于中位数上，那么另一个候选人可以通过在第一个议题上采取与第一个候选人完全一样的立场（从而使该议题变得无关紧要），并在第二个议题上采取中位数立场来赢得选举。同样，如果第一个候选人在第一个问题上取中位数立场但是在第二个问题上却没有，那么另一个候选人可以在两个议题上都取中位数立场来赢得选举。由此可以得出这样一个结论，唯一无法被打败的位置是二维中位数。但是，图20-6却表明，图中用圆圈表示的二维中位数立场是可以被击败的。如果用方形表示的另一个候选人在第一个问题上取偏左的立场，在第二个问题上取偏右的立场，那么就会产生一条分割线，在这条分割线上，他以三票的优势胜出。

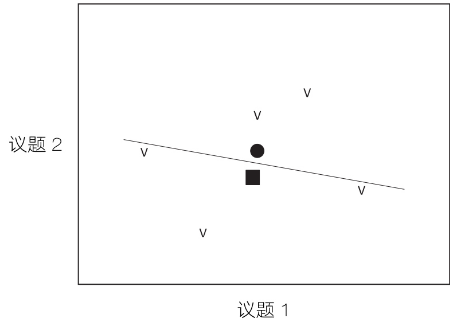


图20-6　二维情况下，中位数候选人会输给挑战者

从这个例子中，我们还可以直观地看出，当且仅当选民理想点的排列，在从中位数出发的每个方向上都不到一半时，二维中位数才可能是不可击败的，这种情况称为径向对称（radial symmetry）。如果选民的理想点在圆形或正方形上均匀分布时，就满足径向对称，这无疑是一个

非常强大的假设。任何位置都是有可能被击败的这个结果，被称为普洛特无赢家结果（Plott's no-winner result），它在两维或多维的情况下都成立。[10]

结果的差异非常明显。一维模型意味着候选人要位于中位数立场，高维模型则意味着他们不应该这么做。我们应该相信哪种模型呢？正确的做法是不应该完全相信任何一种模型，而是从这两种模型中都汲取一些洞见。一维模型揭示了追求选票最大的候选人有趋向温和政策的强大动机，高维模型则证明了这些动机的局限性。没有哪个位置可以保证获胜，所以不应该预期会出现均衡。我们应该期待的是复杂性，观察到各路政客时而联盟，时而拆台，纵横捭阖，翻手为云，无休止地为选票竞争。[11]

我们也可以应用唐斯模型解释美国国会及其各委员会等立法机关和总统制度通过的各种法案的意识形态维度。这里关键是将各个法案映射到与委员会成员相同的单一意识形态维度上。在这里，我们主要考虑三个策略性因素：当前政策的影响（现状效应）、控制议程的权力，以及加入拥有否决权的立法成员的影响。

我们通过一个例子来说明问题。一个由三个人组成的委员会，三位委员的理想点分别为40、60和80，政策提案由理想点为40的委员提出，必须得到多数人同意才能通过。图20-7表明了现状对最终政策的影响。如果现状是80，那么这个提案人就需要中间投票人，也就是理想点为60的那个委员认为该项提案不比现状差。在这种情况下，中间投票人将接受提案人提出的理想点为40的提案，因为它与现状一样好。[12]如果我们将现状移至70，那么这个中间投票人将会拒绝理想点为40的提案，提案者必须提供理想点为50的政策。最后，如果现状位于中间选民的理想点60上，提案者就没有任何权力。因此，我们可以得出以下推论：当现状具有极端值时，提案者拥有最大的权力。

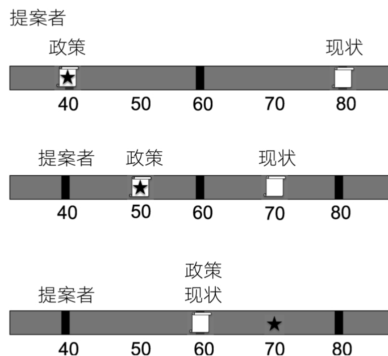


图20-7　现状对最终政策的影响

这个结果适用于需要人们投票做出决策且意见可以映射到一个维度的任何情况下。假设，一个非营利组织的负责人向董事会提出了一个增加用于资助穷人租用经济适用住房的支出建议，如果当前的支出水平与董事会里面的中位数成员的意愿相符，这个负责人的议程权力就非常有限。如果现行政策与董事会成员的理想点不一致，那么他就可以拥有相当可观的权力。

为了说明提出方案的人所拥有的权力，考虑现状为70的情况（如图20-8所示）。在图20-8中，最上面的图显示的是前面说的那种情况，其中提议者的理想点为40，并提出了一个理想点为50的政策。中间的图显示了中位数投票者能够提议自己的理想点60，并且能够获得理想点位于40的委员会成员的支持票的情况。最下面的图则显示了提议者理想点为80的情况。此时，他不能提出他自己和中间选民都更喜欢（甚于现状）的政策，因此只能接受现状。

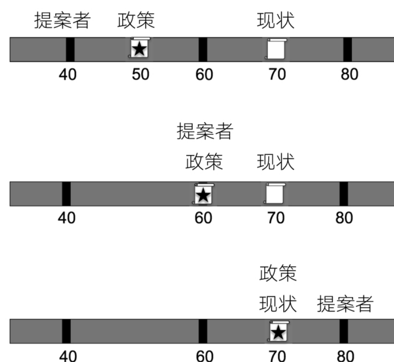


图20-8　提案者对最终政策的影响

这个假想的例子说明了提案者权力的限度。法案可能会朝着拥有权

力的人的理想点的方向靠拢，但是正如在上面已经看到的，对于这种权力会不会过大的担忧，只要考虑一下现状的代表性就可以得到缓解。[13]

最后，我们还可以使用这个模型来讨论多层级政府和更多人拥有否决权的情况。在这里，我们将委员会成员解释为众议院、参议院和总统，并假设要想通过一项法案，必须得到全部这三个投票者的批准。在这种情况下，每一个投票者都拥有否决权。再次参见图20-8中现状为70的情形。如果所有这三个投票者都必须批准才能做出任何改变，那么任何提案都无法击败现状。任何位于70左侧的政策都会被位于80处的投票者否决。另一方面，任何位于80右侧的政策都会被其他两个投票者否决。[14]如果所有这三个投票者都可以否决立法，那么没有任何立法能够通过，除非现状处于40到80之间，也就是说，如果当前现状位于任何人的理想点的左侧或右侧。这个模型揭示了拥有否决权的投票者的数量和意识形态多样性与出现政治僵局的可能性之间的紧密联系。这是一个具有普遍意义的深刻洞见。拥有多样性否决权的组织很可能无法采取任何行动。

享受竞争模型

在享受竞争模型中，各备选项（通常是各种产品）也是用属性来表示的。但是，在这种模型中，属性包括了质量、效率或价格，而且更多或者更少总是更受欢迎的。为了刻画异质性，享受竞争模型允许个体给不同维度赋予不同的权重。

利用线性回归方法，我们可以运用享受竞争模型推断商品属性的隐含价值。具体程序很简单。这种模型假设收益是商品属性和个人赋予权重的线性函数。如果我们获得了数千幢房子的售价以及关于这些房子属性的数据（面积、卧室和浴室的数量，当地学校的质量，院子的大小和建筑质量等），就可以通过产生回归估计出购买房屋的人赋予每个属性的平均权重（以美元计）。这种回归方法通常称为享受回归。

享受竞争模型

备选项包括N个效价属性（valence attribute）：$\vec{a}=(v_1,v_2,\ldots,v_N)$。

个体用各个属性上的权重来表示：$\vec{x}=(w_1,w_2,\ldots,w_N)$。

该个体得自该备选项的收益（效用）为：

$$\pi(\vec{x},\vec{a}) = w_1 \times v_1 + w_2 \times v_2 + \cdots + w_N \times v_N$$

示例：$\vec{v} = (3,1,2)$，$\vec{w} = (4,2,5)$

$$\pi(\vec{v},\vec{w}) = 4\times3 + 2\times1 + 5\times2 = 24$$

其中有些属性，如一个游泳池或新厨房，则是有市场价格的。作为对模型的一种检验方法，我们可以将估计出来的价格与市场成本比较一下。如果回归结果表明，游泳池使房子的价格增加了15万美元，同时建造游泳池的成本却只有15 000美元，我们就知道这个模型遗漏了某些属性。如果回归结果表明增加的价值仅为8 000美元，那么这可能意味着无法收回加建游泳池的全部成本。

其他属性，例如从房屋到市中心的通勤时间，是没有市场价格的。在这种情况下，回归所能得到的是该属性的隐含价格。需要注意的是，隐含价格也可能包括大量信息。表20-1显示了六栋假想房屋的假想价格数据。

表**20-1** 六栋假想房屋的假想价格

房子坐落地	房子面积（平方英尺）	卧室数量（个）	通勤时间（分钟）	价格（美元）
松树街 204 号	2 000	3	30	$200 000
枫树街 312 号	3 000	5	60	$380 000
格迪斯道 211 号	2 500	4	10	$310 000
马丁路 342 号	1 500	2	20	$150 000
克拉克巷 125 号	2 000	4	30	$220 000
布朗苑 918 号	4 000	6	50	$360 000

如果假设房子的所有其他属性都完全相同，那么可以运行享受回归。假设得到的回归方程如下：

价格（美元）＝100（平方英尺数）＋20 000（卧室个数）-2 000（通勤时间）

这就是说，这个回归方程估计，每增加1平方英尺的价值为100美元，每间卧室的价格为20 000美元；另外，通勤时间每增加1分钟影响房子价值2 000美元。假设一个人买下房子后，要住上20年，在20年间，上班的总天数为4 000天至5 000天。如果取较低的那个数字，那么

假设用于通勤的时间每天都能省1分钟，就能省下4 000分钟（或60多个小时）的通勤时间。因此，2 000美元这个估计值相当于每小时30美元左右。换句话说，人们为了住在更接近上班地点的房子里而付出的代价，就“好像”是他们为了免受1小时交通堵塞之苦而付出了30美元一样。[15]

产品竞争的混合模型

空间竞争模型和享受竞争模型在如何表示对各种属性的偏好这一点上有所不同。在空间竞争模型中，每个人对每种属性都有一个更偏好的水平，而且某个备选项在哪些维度上与他的理想点接近，对他来说价值就越大。而在享受竞争模型中，每个属性越多（或越少），人们就越喜欢。

消费品和服务、理想的生活伴侣、公共政策、宗教，以及求职者，全都既包括空间属性，也包括享受属性。每个人也许都有各自喜欢的薯条脆度，但是所有人都希望每份食品的价格更低。脆度是空间属性，价格则是享受属性。不同的雇主在寻找潜在的雇员时，喜欢的人格特征一般会有所不同。有些企业更喜欢外向的人，也有一些企业可能更喜欢内向的人，但是所有企业都更喜欢诚实和正直的人。个性类型是空间属性，而诚实正直则是享受属性。

因此，我们可以创建一个混合模型，让备选项同时包含空间属性和享受属性。这种混合模型可以用来分析市场进入、产品差异化竞争，以及价格竞争的激烈程度。在此，不妨再分析一下前面讨论过的巧克力棒的例子。在确定新产品的属性之前，打算进入巧克力棒市场的企业可能会先将这三种现有产品放在属性空间中分析，然后给消费者发放问卷，以了解他们理想点的分布状况。这个拟进入者可以估计出自己想推出的产品的沃罗诺伊邻域，如果那个邻域的消费者数量很少，就不应期待产品会热销。

任何考虑进入某个市场的企业家都可以采用这种方法。一个靴子设计师可以绘制出现在所有绝缘靴的设计图（那很可能有数十种），然后他可能会发现没有一种设计是用闪亮漆皮的。设计一个用于制作待办事项列表的智能手机应用程序的开发者，可以列出现有类似应用程序的功能，估计市场总需求并预测可能的销售额。

我们可以将空间竞争模型中的降价行为视为分割线的移动。回顾一

下前面的图20-3，图中显示了两种巧克力棒的情形。分割线对应于那些在A巧克力棒与B巧克力棒之间无差异的消费者的理想点。如果生产B巧克力棒的企业进行了降价促销，并且如果消费者更愿意付出更少的钱来得到同样多的糖果，那么这种降价会把分割线推向A并使B的市场份额增大。我们不需要模型就知道B的降价会增大其市场份额，但是我们确实需要这个模型来估计这种影响的大小。关键在于要将拥挤的市场（crowed market）与稀疏的市场（sparse market）区分开来。前者指在一个低维的属性空间挤进了大量的产品，后者则指市场中只有很少的竞争对手。在拥挤的市场中，每种产品都只有一个小小的沃罗诺伊邻域，而在稀疏的市场中，沃罗诺伊邻域是很大的。

价格变化在这两种不同类型的市场中会产生不同的影响。图20-9显示了B巧克力降低10%（假设价格从每根2美元下降为1.8美元）的影响。左图显示的是一个稀疏的市场中的情形。降低B巧克力棒的价格会改变A与B之间的分割线，并将B的市场份额从50%提高到54%（B的市场份额增加了8%）。价格下降了10%，但是销量只增长了8%，综合起来使销售收入减少了3%。因此，在这种情形下，降低价格不是一个好主意。右图显示的是一个拥挤的市场中的情形，市场中有7种不同的巧克力棒。在这里，价格下降对B的绝对市场份额影响也不算太大，只是从15%增加到了20%，但这5%的增幅却代表着B的市场份额有相当大的增长比例（达到了33%）。总体影响是使销售收入增加了20%。[16]

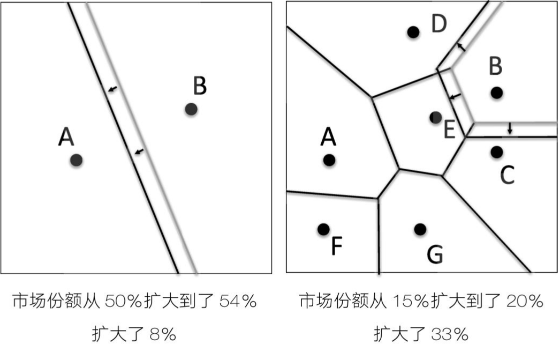


图20-9　稀疏的市场与拥挤的市场中的价格竞争

因此，这个模型的预测是，拥挤的市场中的价格竞争比稀疏的市场中的价格竞争影响更大；同时，大宗商品(7)市场上的价格竞争则可以称得上是“极端的竞争”，其影响最大。这个模型还预测，高端时尚商品的价格竞争将很少出现，因为产品的高维性能够创造出一个稀疏的市场，设计师可以维持大幅度的价格加成。

属性数量与价格竞争激烈程度之间的这种关系表明，增加新属性不失为一个好策略。这将使市场更加稀疏，从而降低价格竞争烈度，并带来更高的利润。然而，即便这种推断本身是正确的，要实现这个策略也并不容易。无论如何，人们必定会给新的属性较高的估值。事实上，每一个成功的新属性的出现（例如无绳立体声扬声器），背后都隐藏着多次失败尝试的历史（例如比克公司命运多舛的一次性产品）。

小结

空间竞争模型、享受竞争模型，以及两者的混合模型提供了一个很好的框架，在这个框架内，我们可以表征不同的产品、政治候选人，甚至是求职者。这些模型可以测量意识形态“位置”、隐含价格属性，并评估潜在的市场进入者。它们提供了很多深刻的洞见，有助于我们理解市场竞争如何产生差异化的激励、政治竞争如何产生建立联盟的激励，以及只具有较少属性的产品的价格竞争为什么会更加激烈。

这些模型都做出了相当强的、从经验上看相当可疑的假设。例如，在前面讨论的模型中，我们假设人们的偏好是不会改变的，他们不会屈服于社会压力。如果真的是这样的话，为什么企业和政客会花费大量金钱和资源去改变消费者和选民的偏好呢？对于这种批评，我们可以再次援引博克斯的格言来辩解：所有的模型都是错的。

还有一种更精细的回应方法可以对基本偏好与工具偏好加以区分。前者指人们的欲望（想得到什么），后者指人们对能够产生根本性结果的属性的偏好。学生的基本偏好可能会在成为一个受欢迎的人、身体健康和取得学术成就之间维持平衡。为此，一个学生可能会通过提早起床、喝果汁、尽快完成自己的家庭作业（以保证晚上有时间参加社交活动），来实现这些根本性的目标。而这个学生对水果奶昔的选择则是工具性的，水果奶昔有助于实现他对健康的基本偏好。如果哪一天，他看到一篇科学论文说水果奶昔的糖分很高，不利于身体健康，他就可能会改喝饮用水。如果是这样，他的工具偏好就会发生改变，但是他的基本偏好并没有变化。在这里，我们再一次看到，模型本身并不是目的，只是提供了一个框架来构建我们的思维。

多个价值模型

在市场经济中，我们可以通过一个愿意支付的价格来衡量一件商品对这个人的价值。一个人可能会认为一个五香熏牛肉三明治值7美元、戈雅（Goya）[8]的一幅画值300万美元、佛罗里达州奥卡拉市的一块土地值75 000美元。但是，许多经济模型都忽略了这些估值的来源。美国经济学家乔治·施蒂格勒（George Stigler）引用了大文豪陀思妥耶夫斯基（Dostoyevsky）的《卡拉马佐夫兄弟》（The *Brothers Karamazov*）一书中那个著名的情感主义者米蒂亚（Mitya）说的一句话。他说："De gustibus non est disputandum。"（口味各异，难言好坏）我们这里所讨论的模型确实比较少讨论口味本身，而更多讨论口味如何转化为人们赋予商品的货币价值。

享受模型：这个模型根据商品的内在属性解释商品的价值。估值的差异取决于对商品属性的不同潜在偏好。这些属性可能是商品的物理组成部分。例如，一个五香熏牛肉三明治中包括了黑麦面包、五香熏牛肉、瑞士奶酪、芥末、泡菜和洋葱。然后，我们可以将它的价值写成这些组成部分的价值的加权线性组合。更精细的享受模型还可以包括交互项。例如，如果与烤黑麦面包一起售卖，那么五香熏牛肉可能会更有价值。

协调模型：这种模型将价格解释为社会建构的。戈雅画作的价值取决于其他人相信它真的有那么高的价值。在最初，人们对戈雅画作的价值有自己的看法。然后，他们与社会网络中的其他人互动并更新自己的看法。两个人都可以将他们的价值设定为等于两人估价的平均值，一个人可以改变他的估价以匹配另一个人给出的价值，或者每个人都可以根据其他人的估价来调整自己的估价。给定这三种假设中的任何一种，估价都可以实现局部收敛。这样，相互之间有联系的人将会具有类似的估值。分配给商品的最终价值将取决于价值的初始分布、社会网络以及配对发生的顺序。

预测模型：这类模型将价格解释为对未来价值的预测。根据这种模型，佛罗里达州奥卡拉市的土地价值、比特币或股票价值，都取决于人们将来愿意为它们付出多少钱。这些估值取决于预测模型，而预测模型又取决于属性和类别。我们可能会将奥卡拉市归类为气候温暖的、税率很低的非沿海城市。人们的估值的变化源于不同的预测模型。投资者会使用多种预测模型。这些模型可能依赖于属性，或者，就像对比特币进行估价一样，也会对协调做出假设。

这三种模型为商品的价值提供了三种不同的解释。没有任何一个模型在所有情况下都是最好的。在特定情况下，每个模型都可能成为最好的那个。模型就像箭袋中的箭。五香熏牛肉三明治的价值很可能来自其内在性质，戈雅画作的价值则可能在很大程度上是社会建构的——只要人们认为他的画作有价值，那么它们就有价值。对佛罗里达州土地价格的估计则可能取决于对未来房地产价值的预测。

21　博弈论模型

演绎推理是从最抽象到最不抽象的推理。它从一套公理开始，运用逻辑定律和数学规律来操纵，形成对世界的预测。

雷切尔·克罗松（Rachel Croson）

本章以及随后各章中讨论的许多模型，包括合作模型、信号模型、与机制设计有关的模型和与集体行动有关的模型都要涉及博弈。我们不会在这里非常深入地探讨博弈本身，因为实际上整本书都是讨论这个主题的。本章的目标是提供一个适当的入门介绍，为此，我们给出了三种主要类型的博弈的一些例子：标准式博弈，博弈参与者在一组离散的行动（通常为两种）中做出选择；序贯博弈（sequentail game），博弈参与者按顺序选择行动；连续行动博弈，博弈参与者可以选择任意尺度或效果的行动。我们通过这些例子介绍了博弈的主要概念，有助于理解后面章节中给出的模型。当然，它们本身也有讨论的价值。

在本章的其余部分，我们首先讨论2×2标准式零和博弈。在零和博弈中，两个博弈参与者中的每一个都要在两个行动中做出选择，无论某个博弈参与者选择什么行动，一个博弈参与者得到的收益，都会被另一个博弈参与者遭受的损失所抵消。我们利用零和博弈的例子来定义博弈论的基本术语，区分策略和行动，并引入迭代消除被占优策略的概念。然后，研究市场进入博弈（market entry game），市场进入博弈是序贯博弈的一种。我们将在重复市场进入博弈的框架下讨论所谓的连锁店悖论（chain store paradox）。接着，考虑一个“努力博弈”，这是一种连续行动博弈。在这个博弈中，个体选择努力水平以赢得固定金额的奖励，付出努力越大，博弈参与者赢得奖励的机会越高。本章最后简要讨论了博弈论模型的一般价值。

标准式零和博弈

在本节中，我们分析两个双人标准式零和博弈（two-player normal-form zero-sum games）。在这种博弈中，每个博弈参与者选择一个行动，并根据博弈参与者自己的行动和另一个博弈参与者的行动获得一定收益。此外，博弈参与者双方的收益总和为零。

在图21-1所示的第一个硬币配对博弈中，每个博弈参与者都要在两个行动中做出选择：猜硬币正面朝上还是背面朝上。行博弈参与者希望自己的选择与另一个博弈参与者的选择匹配，而列博弈参与者则不希望匹配。收益如矩阵所示：

	正面朝上	背面朝上
正面朝上	**1**，−1	**−1**，1
背面朝上	**−1**，1	**1**，−1

图**21-1**　硬币配对博弈

博弈的策略是如何进行博弈的规则，它可以是对单个动作的选择、在不同行动之间的随机化，或者，正如在下一节中将会看到的，也可以是一个行动序列。博弈的纳什均衡（Nash　equilibrium）是指这样一种策略，它们能够使每个博弈参与者的策略在给定其他博弈参与者策略的情况下是最优的。在硬币配对博弈中，存在一个唯一的均衡策略，那就是，两个博弈参与者都以相同的概率在两个行动之间进行随机化。为了证明随机化是一种均衡，只需要证明，如果某个博弈参与者随机化，那么另一个博弈参与者选择任何行动都不可能比随机化更好。

要证明这一点很简单。如果行博弈参与者（在图21-1中，行博弈参与者的行动以黑体显示）以1/2的概率选择正面朝上、1/2的概率选择背面朝上，无论他的选择到底是什么，列博弈参与者的收益都为零。正因为如此，随机化是列博弈参与者的最优策略。根据对称性，随机化也是行博弈参与者的最佳选择。

随机化策略的最优性，对策略互动环境中的行为有很大的意义。体育运动也是零和博弈：一方获胜，另一方就要落败。在点球大战中，一名前锋希望在瞄准球门左侧与球门右侧之间进行随机选择；在网球比赛中，发球方要随机将球发到内角或外角；在足球比赛中，进攻方希望在跑动与传球之间随机选择。而且，在所有这些比赛中，另一方也会随机化他们的反应。任何非随机性都可能会被对手利用，扑克等纸牌游戏也是如此。一个优秀的扑克玩家会随机地虚张声势。如果他一直虚张声势，对手就会了解这种策略，他就会落败。当然，对手的最优策略也是随机地虚张声势，这样就同样有可能赢或输。

现在讨论最小化风险博弈（minimize risk game）。在图21-2所示的这个博弈中，每一个博弈参与者都可以选择采取冒险的行动或安全的行动，这是一个非对称的零和博弈。博弈参与者的收益不仅取决于自己的行动，还取决于哪一个博弈参与者采取了哪一个行动。在这个博弈中，行博弈参与者有一个占优策略，即采取安全的行动。无论列博弈参与者选择哪一个种动作，对于行博弈参与者来说，选择安全的行动总是更好的。但是对于列博弈参与者来说，情况却并非如此。如果行博弈参与者选择冒风险，那么列博弈参与者也应该选择冒风险；如果行博弈参与者选择了安全的行动，那么列博弈参与者也应该选择安全的行动。

	选择冒险	选择安全
选择冒险	−10，10	0，0
选择安全	10，−10	0，0

图21-2　最小化风险博弈

通过考虑对行博弈参与者激励的情况下，列博弈参与者可以推断出行博弈参与者总是会选择安全的行动，那么他也会选择安全的行动。这种一方为另一方排除最优策略被称为迭代消除被占优策略。因此，两个博弈参与者都选择安全的行为是这个博弈的纳什均衡。

序贯博弈

在序贯博弈中，博弈参与者按照某个特定的顺序采取行动。由此，可以用一棵博弈树（game tree）来表示一个序贯博弈。博弈树由节点和边组成，每个节点对应于博弈参与者必须采取行动的时刻，该节点的每条边分别表示可以采取的某个行动。在博弈树最末尾的分支上，我们写下相应行动路径的收益。图21-3所示的博弈树显示了市场进入博弈。

在市场进入博弈中，有两个博弈参与者：拟进入者和现有企业。如果拟进入者选择不进入市场（博弈树的左侧分支），那么它的收益为零，现有企业的收益为5。如果拟进入者决定进入市场，那么现有企业必须做出选择：是接受新进入者，同时自己的收益从5下降为2，还是发动与新进入者的商战，但这会导致自己的收益变为零，同时令新进入者

的收益为负。之所以假设这种情况下新进入者的收益为负，因为它必须为进入市场付出一定的成本。

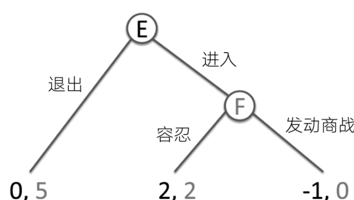


图21-3　市场进入博弈

在序贯博弈中，策略对应于每个节点处的行动选择。假设现有企业在发现有新企业进入时决定发动商战。那么，如果拟进入者知道这一点，就不会选择进入，因为这种情况下进入会产生负收益。这个行动序列——拟进入者选择不进入、现有企业在拟进入者进入时就会发动商战，是一个纳什均衡。然而，这并不是唯一的纳什均衡，也不是最有可能出现的结果。拟进入者选择进入市场，现有企业决定接受（不发动商战），这是第二个均衡。

那么，应该如何在这两个均衡之间做出选择呢？我们可以利用细化准则。在序贯博弈中，一种常见的细化准则是选择子博弈完美均衡（subgame perfect equilibrium）。可以运用逆向归纳法（backward induction）来求解子博弈均衡：从最末端的节点开始，并在每个节点处选择最优行动。然后沿着博弈树逆向倒推，假设每个博弈参与者会在给定另一个博弈参与者在后续节点上的行动时选择最优行动。例如，在市场进入博弈中，我们从现有企业的末端节点开始推导。它有一个最优行动，即接受对方进入。然后移动到博弈树上面的节点，不难发现拟进入者的最优策略是进入。

这个博弈在重复进行时会变得更加有趣。试想一下，现有企业也可能存在于许多个市场中。也许它是一家连锁企业，在几十个城市都有门店。再假设存在一系列的拟进入者。那么，这个企业将陆续地进行一系列市场进入博弈。

如果现有企业从最后一个市场开始使用逆向归纳法进行推理，那么它将接受最后一个市场中的进入者。根据同样的逻辑继续推导，现有企业将接受倒数第二个市场中的进入者，以此类推，它将接受所有市场中的进入者。在序贯博弈唯一的子博弈完美均衡中，所有潜在的进入者都选择进入，现有公司接受所有。

虽然每个市场的进入和接受都是唯一的子博弈完美均衡，但实际上，这可能不会发生。假设我们是现有公司董事会的成员，我们面对的是第一个进入者曾经学过博弈论并已经进入市场。我们可能想要竞争，试图阻止其进入其他市场。如果竞争是可信的，那这将是一个明智的策略，也就是说，如果能够建立一个愿意竞争的声誉。我们希望创造的结果不同于子博弈完全均衡。

博弈理论家将这种情况称为连锁店悖论。这是一个例子，博弈论认为的最优行为可能不是一个老练的行为者在利害关系很大时所选择的行为。这个例子并没有反驳博弈论或破坏理性选择假设，而是揭示了为什么我们总是必须挑战假设。

连续行动博弈

我们现在研究另一种博弈。在这种博弈中，博弈参与者可以在连续的可能行动集中进行选择。在连续行动博弈中，行动对应于努力水平。通过选择更大的努力，博弈参与者能够增大自己赢得奖励的概率。这个博弈还允许考虑任意大数量的博弈参与者。

努力博弈

N个博弈参与者中的每一个人都要选择以货币形式表达努力水平，以赢得价值为M的奖励。一个博弈参与者赢得奖励的概率等于他的努力水平除以所有博弈参与者的总努力水平。如果令E_i表示博弈参与者i的努力水平，那么他的获胜概率由以下方程式表示：[1]

$$P(i\ wins) = \frac{E_i}{(E_1 + E_2 + \ldots + E_N)}$$

均衡努力水平为：$E_i = \frac{M}{N} - \frac{M}{N^2}$

均衡努力水平的表达式揭示了很多重要的含义，正如我们所预料的那样，个人的努力水平会随着奖金的增多而增大。同样，在均衡状态下，总努力水平将会小于奖金的价值。在假设博弈参与者会进行最优化的情况下也会得到这些结果。博弈参与者应该付出一定努力以赢得奖励，但是不应该付出不合理的努力水平。

通过增加博弈参与者的数量，可以看到其对个人和总体努力的影

响。根据模型，即便每个人的努力水平都下降了，总努力水平也会增加。这个结果说明，那种吸引了大量“参赛者”的研究课题竞标、建筑设计竞赛和征文比赛，反而可能会产生水平不那么优异的赢家（与“参赛者”较小的竞赛相比），因为在这种“参赛者”众多的竞赛中，个体参与者付出的努力水平会较低。

小结

本章一开始，我们先讨论了零和博弈，这类博弈不包括互利的行动组合。任何对一个博弈参与者有益的行动都必定会损害另一个博弈参与者。在零和博弈中，任何行为对一方的“益”，在数量上都等于对另一方的“损”。从一个人那里拿钱给另一个人，就是一种零和行动。许多个人行动和政策选择至少在一个方面是零和的，我们每天只有这么多的时间可用，这么多钱可花，这么多资源可分配。也就是说，在这个维度上的零和行动，在另一个维度上可能不是零和的。例如，重新安排预算，在货币的数量这个维度上是零和的，但是就人的幸福感或满足程度这个维度而言，却可能是正和的或负和的。

我们应该始终探究提议的政策变化是否会导致零和博弈。许多人都对家长的择校行为颇有微词（这里所说的择校问题，主要是父母将孩子送进什么样的学校学习的能力），因为它加剧了竞争。但是迫于竞争的压力，学校的教学质量将会得到提高，至少从逻辑上看是这样。

然而，只有在产能过剩的情况下，学校才有动力提高质量，否则，择校会在学生中造成零和博弈。假设一个城市有1万名学生和10所学校，每个学校可以接受1 000名学生。如果所有学生都以相同的方式对学校进行排名，那么最好的学校的名额将只能通过抽签来分配，中签的学生将去更好的学校，未中的学生将去更差的学校。学生们玩的是一场零和博弈。如果新的学校开放或现有学校改善，就将不再是零和博弈，每个人都可以赢。

市场进入博弈与零和博弈都提供了有益的洞见。市场模型揭示了学校改进质量的动力。零和博弈则表明有些学生将受益、有些学生将受损。每个人所承受损害或获得的收益的大小，则取决于具体情况：学生和家长对学校的质量的了解程度如何，学校还有存在多少剩余名额，学校是否真的知道如何提高教学质量，会创办新的学校吗？

这两个博弈都没有给我们一个正确的答案，但是都产生了有用的见

解。择校会带来竞争。它还产生了一个具有零和博弈特征的大规模排序问题。竞争的积极方面是否会超过消极的排序成本取决于环境。我们必须把模型排列在一系列事实的基础上，才能做出正确的政策选择。

识别问题

关于人们行为的数据经常揭示出人类行为的“聚类倾向”。优秀的学生更有可能与其他优秀的学生成为朋友，而不怎么可能与学习困难的学生成为朋友。有过犯罪前科的人比从未犯过罪的人更有可能与犯罪的人交往。在社交网络中，各种各样的“社会善”和“社会恶”——吸烟嗜好、健美的人、肥胖者，甚至幸福也都会“聚类”。人们还会根据信仰或意识形态聚在一起。

有两个模型可以解释这种聚类：同伴效应模型（peer effect model）和分类模型（sorting model）。同伴效应模型用博弈论来解释聚类现象，即一个人与他的朋友一起进行协调博弈。而在分类模型中，人们会“迁移”到与他们相似的其他人附近。一群优秀的学生之所以聚在一起，可能是因他们要协调完成某个共同行动（同伴效应），或者也可能是因为优秀的学生就喜欢找优秀的学生一起玩（分类效应）。如果只有数据快照（snapshot of data），那么这两者是无法区分的。

数据：学生既有可能获得高分H，也有可能获得中等分数M，两者是等可能的（概率相同）。假设每一个学生都会加入某个人数为4的“小团体”，则有如下分布：p（{H，H，H，H}）＝P（{M，M，M，M}）＝5/16，p（{H，H，H，M}）＝P（{M，M，M，H}）＝3/16，p（{H，H，M，M}）＝0。

同伴效应模型：学生最初先会形成一些人数为4的随机小组：p（{H，H，H，H}）＝P（{M，M，M，M}）＝1/16，p（{H，H，H，M}）＝P（{M，M，M，H}）＝1/4，并且p（{H，H，M，M}）＝6/16。属于仅包含一种类型的小组的人数保持不变。与小组内所有其他人类型都相反的人会切换类型，因此，{H，H，H，M}这样的小组会变为{H，H，H，H}。在每种类型的人数相同数量的小组中，会有一个成员切换类型，也就是说，{H，H，M，M}这样的小组有同样的概率会成为{H，H，H，M}或{M，M，M，H}。

分类模型：学生最初先会形成一些人数为4的随机小组。在具有两种类型的任何一个小组中，与至少两个其他人的类型相反的人，会与相反类型的某个人交换小组。也就是说，{H，H，H，M}会变为{H，H，H，H}，而且{M，M，M，H}会变为{M，M，M，M}；并且，任何{H，H，M，M}这种形式的小组有同样的概率成为{H，H，H，M}或{M，M，M，H}。

两个模型都与数据一致，从而导致了识别问题。只有数据快照，我们无法确定吸烟、看漫画书或喜欢滑板到底因为同伴效应导致的，还是因为分类效应导致的。在某些情况下，我们可以推断出使用哪种模型更好。例如，美国中西部地区的人们喜欢说“pop”（泡泡）、沿海地区的人们喜欢说“soda”（苏打），对于这种倾向，我们可以有把握地认为这是由同伴效应驱动的，因为很少从外地移居波士顿的人，会将可口可乐称为“苏打水”。但是，对于某些利害关系更大的行为，例如学业成绩、滥用药物、肥胖和幸福，就需要用更多的时间序列数据来识别哪种模型才适用了。利用时间序列数据，我们就可以分辨出人们到底是在改变自己的行为（同伴效应），还是在更换他们的朋友（分类效应）。在许多情况下，这两个因素都有。[2]

22　合作模型

从来没有人因施舍而变得贫穷。

安妮·弗兰克（Anne Frank）

经常会有人要求科学家列出他们心目中最重要的问题，各个领域的专家给出的最重要的问题往往各不相同，例如，宇宙是如何形成的？意识是怎样出现的？我们能否找到根治癌症的方法？等等。在这些问题当中，有一个是社会科学家和生物科学家都认为是重要的问题，那就是，合作是怎么产生的？[1]合作要求合作者采取不符合自身利益的行动，而这就意味着我们不会经常观察到合作现象。但是在现实世界中，我们却看到合作出现在无数领域中，而且达到了非常大的规模。合作在细胞层面上就存在：细胞通过黏附作用实现了合作，一个细胞会产生细胞外物质，供其他细胞黏附之用。我们观察到，蚂蚁、蜜蜂、人类、人类组织之间，甚至国家之间都存在着广泛合作，不同国家会在制定条约和国际法方面进行合作。

在本章中，我们运用模型来讨论合作如何产生、如何维持以及怎样做才能创造更多的合作。当然，本章给出的这些模型无法完美地解释世界上广泛存在的各种合作，比如，为什么乌鸦会告诉其他乌鸦所发现的腐肉地点，为什么裸鼹鼠会共同防御以它们为猎物的天敌，为什么攀爬藤蔓种植在“亲属”旁边时根系较不发达，为什么白蚁和蜜蜂会建造出精致复杂的巢穴，为什么蚂蚁能用身体和附肢搭起运送食物的桥梁，等等。但是，这些模型确实能够告诉我们很多重要的结论。[2]

尽管我们看到，物种内部和物种之间广泛存在着合作的例子，但是我们同时还应该看到合作失败的情况。合作的程度取决于具体环境。联盟既能吸引人们加入，也会失去原有成员；英国是欧盟的创始成员国之一，但是后来却要退出欧盟。一个积极为学校筹款活动服务的人，却可能在超市排队时插队或者偷税漏税。一头狮子，既可能加入围猎水牛的团体，也可能自己偷偷地去猎杀野猪。并不是每个物种都合作。黑胡桃树的根须将一种名为胡桃醌（juglone）的物质释放到土壤中，以抑制附近植物的生长。

细胞、树根、乌鸦、人、企业和国家等合作型实体行为特征的多样性，要求采取多模型方法。也许，我们最好将细胞和植物建模为遵循固

定规则的，将乌鸦、蚂蚁和狮子建模为运用更多依赖于环境或过去结果规则的，将人、企业和国家建模为有前瞻能力并会进行成本收益计算的。

本章的第一个要点是：合作可以通过多种机制、在多种环境下涌现出来并维持下去。我们讨论了四种促进合作的机制：重复、声誉、局部聚类和群体选择。这些机制都能在没有外部干预或管理的情况下进行合作。它们可以适用于合作的鼹鼠、蜜蜂和人类，人类还有更正式的促进和保持合作的方式。在本章最后的讨论中，我们描述了其他制度性的解决方案，包括付钱让人们进行合作，如果不这样做就惩罚，或者制定法律强制人们的合作行为。

本章的第二个要点是，这些机制中任何一个的“效力”都取决于那些正在合作的实体所拥有的“行为曲目”（behavioral repertoires）。一些机制，特别是通过重复实现合作这个机制，几乎适用于任何行为。声誉和规范这两个机制则需要前瞻性行为和信息共享，对于那些更“老练”的行为者来说更加有效。

局部聚类对合作的效果取决于具体模型。由演化力量选择支持或反对的行为者之间的合作，最常出现在稀疏的网络上，而通过规范进行合作则需要密集的网络。群体选择的有效性则取决于行动者的前瞻能力和适应速度的细微差别。行为者所拥有的更强的前瞻能力，能够增强群体选择的力量，而允许行为者更快地适应则可能会阻碍群体选择发挥作用。为了探索这些问题并分析清楚行为假设与合作结果之间的相互关系，要利用我们熟悉的囚徒困境博弈模型和合作行动模型。合作行为模型允许我们刻画有利于多个参与者的行动，并对网络上的合作行为进行建模。

本章安排如下。我们首先描述了囚徒困境博弈，同时说明理性行为者之间的合作是怎样得以维持的。接着，我们阐述了重复行动是如何促进基于规则的行动者之间的合作的，以及为什么不断发展合作比单纯维持合作更加困难。然后，我们考虑不太复杂的生物学意义上的行为者之间的合作，并阐明了亲缘选择和局部聚类如何促进了合作。最后两部分讨论了群体选择以及我们如何综合这些模型实现更高程度上的合作的问题。

囚徒困境博弈

囚徒困境博弈的名称源于如下故事。有两个人，被控共同犯下了某种罪行。有关当局只掌握了间接证据，因此给他们每个人都提供了认罪减刑的机会。两人因此面临着两难选择：如果两人都不认罪，那么每个人都会（根据现有证据）受到轻微的惩罚；如果只有一人认罪，那么认罪的这个人不会受到惩罚，而另一个人则会受到很严厉的惩罚；如果两人都认罪，那么两人都会受到较严厉的惩罚，但是不会像只有一个人认罪时那么严厉。

图22-1将这个故事表述为一个双人博弈。每个博弈参与者可以选择合作或背叛。图中的灰色数字表示列博弈参与者的收益，黑色数字表示行博弈参与者的收益。对每个博弈参与者来说，背叛都是占优策略，无论其他博弈参与者采取什么行动，背叛都能带来更高的收益。但是，如果两个博弈参与者都背叛，每个博弈参与者的收益都会低于双方合作时的收益。因此，追求自身利益的行为导致了集体利益的恶化。

	合作	背叛
合作	3，3	1，4
背叛	4，1	2，2

图22-1　囚徒困境博弈的例子

囚徒困境博弈紧紧抓住了现实世界中许多情况下的核心激励，它可以用来建模一些国家之间的军备竞赛：背叛对应于将资源用于开发武器，合作对应于发展经济。还可以用来建模竞选活动中的广告战：背叛对应于投放负面广告，合作对应于投放正面广告。它甚至可以用来解释为什么雄孔雀会有如此之长的尾巴，每只孔雀都有很强的动力使自己看上去比其他孔雀更强壮、更健美。

很多囚徒困境博弈都是在事后才认识到的。许多新技术的最早一批采用者，例如，最早使用ATM机的银行，发现自己的利润因此大为增加。但是，当其他银行也跟进时，利润就因竞争加剧而下降了。事后证明，可以把选择使用ATM机类视为一种“背叛”。[3]

如图22-2所示，一般形式的囚徒困境博弈假设如果两个参与者都选择背叛，那么基线收益为零。这样一来，我们就可以用三个变量来表示这个博弈：来自合作行为的奖励R，背叛的诱惑T，以及被损害的一方的

收益S（参见图22-2中的收益矩阵）。收益矩阵下面的不等式确保了选择背叛是一个占优策略，而选择合作则能够产生有效率的结果。

	合作	背叛
合作	R，R	–S，T
背叛	T，–S	0，0

T > R > 0 > – S
2R > T – S

图22-2　一般形式的囚徒困境博弈

通过重复和声誉机制实现合作

我们首先阐明，博弈的重复进行和声誉的建立为什么能够维持理性行为者之间的合作。能够维持合作这个事实，并不能保证合作真的能够实现，它只是说明，如果合作“不知怎么”出现时，理性的博弈参与者能够维持它。为了证明重复行动能够维持合作，我们构建了一个重复博弈模型。在这个重复博弈中，每次博弈结束之后，都以概率P再一次进行。从理论上说，这种博弈可以永远持续下去。

博弈参与者在重复博弈中，要根据以往的博弈历史选择行动。在这里，我们考虑一个被称为冷酷触发（grim　Trigger）的重复博弈策略。具体来说，这个策略是，在第一次博弈中选择合作，并且，只要另一个博弈参与者不背叛，那么就在未来的所有博弈中一直选择合作；但是，一旦另一个博弈参与者背叛了，那么就永远选择背叛。冷酷触发策略是“永不饶恕”的。如果两个博弈参与者都采用冷酷触发策略，那么双方将会永远合作。

要想证明冷酷触发策略能够在重复博弈中维持合作，我们只需要证明，如果一个博弈参与者选择了冷酷触发策略，那么另一个博弈参与者也可以通过采用冷酷触发策略获得最高的收益。由于第二个博弈参与者的任何一个偏离合作的行为都会导致第一个博弈参与者无休止的背叛，所以第二个博弈参与者只需要对一直合作的预期收益，与一次背叛再加上两个博弈参与者此后都一直背叛的预期收益加以比较就行了。[4]而冷酷触发策略能否带来更高的收益，则取决于诱惑的大小、合作回报的多少以及重复博弈的概率。

重复博弈维持合作

在重复囚徒困境博弈中，如果继续进行下一次博弈的概率P超过了诱惑收益T减去奖励收益R的差与诱惑支付的比，那么采用冷酷触发策略就能维持合作，即：[5]

$$P > \frac{(T-R)}{T}$$

这个结果告诉我们，如果诱惑收益超过了奖励收益的三倍，即$T>3R$，那么博弈必须以超过2/3的概率重复。这个不等式还告诉我们，如果合作的奖励增大了、博弈重复进行的可能性增加了，或者背叛的诱惑减少了，合作就会变得更加容易维持。这些含义中的每一个，都意味着一条直观的促进合作的途径：增加合作奖励，让重复进行博弈的可能性更大，以及减少背叛的诱惑。虽然这些都只是非常简单的推论，但是在写出这个模型之前，它们可能并不是一目了然的。

通过思考合作的必要条件，我们还可以推断出一些不那么直接的结论。上面这个不等式还意味着，如果博弈参与者认为博弈重复进行的概率在未来会下降到低于阈值，那么理性的博弈参与者将会在这种概率变化发生之前就停止合作，而不会等到变化发生时再停止合作。[6]

而且，重复博弈能不能维持理性的博弈参与者之间的合作，还取决于这个模型的一个特别假设：博弈会以一定概率不断重复进行下去。相反，如果转而假设博弈只会重复一定次数，比如只会重复进行3次，那么理性的博弈参与者将不会选择合作。

我们可以利用逆向归纳法来证明这一点。假设博弈只重复进行3次并且第一个博弈参与者将采用冷酷触发策略。再假设$T=3$、$R=2$、$S=1$。给定这样的收益矩阵，如果第二个博弈参与者在所有三轮博弈中都选择合作，那么他的总收益为6。接下来需要确定是不是所有其他策略都不能带来更高的收益。如果第二个博弈参与者在第一轮博弈中就背叛，那么他所能得到的收益仅为3，因为在他背叛之后，第一个博弈参与者将在后面两轮博弈中都选择背叛。如果第二个博弈参与者在第二轮博弈时再背叛，那么他可以得到的收益为5。但是这两个策略都是不理性的。如果在第三轮博弈中才背叛，那么他可以得到的收益为7（前两轮博弈中各得到2，最后一轮博弈中得到3）。因此，理性的博弈参与者会在最后一轮博弈背叛。

但是，第一个博弈参与者（他宣布自己会采取冷酷触发策略）应该

会意识到第二个博弈参与者会在第三轮博弈中背叛，因而他也会在第三轮背叛。第二个博弈参与者也会意识到两个博弈参与者都会在第三轮博弈中背叛，因此他在第二轮博弈中就会背叛。根据同样的逻辑，第一个博弈参与者也会在第二轮博弈中就背叛，这种推理也适用于第一轮博弈。事实上，只要博弈只重复有限次数（无论多少次），上述结果就都适用。在最后一轮博弈中，理性的博弈参与者会背叛。结果，两个博弈参与者都有动力在倒数第二轮博弈中背叛，以此类推，他们会在所有轮次的博弈中都背叛。从而，唯一符合理性的策略就是永远背叛。

目前为止，我们一直是在一个孤立的环境中考虑两人博弈的，并没有考虑一个人的背叛行为可能会影响其他人对待背叛者的态度。这样做，其实是将这种两人博弈圈进了一个封闭世界中。我们可以扩展这个模型，让博弈在一个社区的成员之间进行，并让这些人有机会监督其他参与博弈的人的行为。

现在假设，每一天，先让这些人随机两两配对，然后进行囚徒困境博弈。在这种情况下，我们假设这个社区的所有成员相信博弈将会永远重复进行下去，因此未来继续博弈的概率等于1。在这些假设条件下，每一个人都不太可能在第二天仍然与前一天博弈过的那个人博弈，所以他们背叛的动机会更加强大。但是，由于我们假设同一社区的人可能会认出谁是背叛者，所以那些背叛过的人很可能会留下不好的名声，根据假设，这个社区中的任何人将来都不会与声誉不好的人合作。

如果我们用P_D表示一个人背叛且被发现、从而留下了一个背叛者的坏名声并在未来的所有博弈中受到惩罚的概率，那么，通过声誉机制维持合作的条件，也就与重复博弈条件下维持合作的条件一样了。这个条件可以写为$P > \frac{(T-R)}{T}$，用一个背叛被发现的概率取代博弈重复进行下去的概率P。

在声誉模型中，合作是通过社区来实施的。背叛并被发现的人，在未来将会遭到所有博弈参与者的背叛。在这里，个体还是会计算背叛的收益和成本。他们还必须相信其他人会坚持惩罚到底，而这也就意味着其他所有人都会背叛那些背叛的人。要做到这一点，所有个体必须要么彼此认识，要么有某种方法来识别或标记过去的背叛者。因此，在其他条件相同的情况下，规模较小的社区的成员应该能够更好地通过这种机制来实施合作。在美国北方的小城镇，人们在冬季会将车停在商店停车场里。他们不用担心汽车会不会被人偷走（“背叛”），因为他们认识城里的每个人。任何偷车的人，即便是恶作剧，都会导致自己的声誉下

降。

实物标签可以使声誉变成一个公共信息，从而有利于能够维持合作。在纳撒尼尔·霍桑（Nathaniel Hawthorne）的小说《红字》（*The Scarlet Letter*）中，海丝特·白兰（Hester Prynne）被迫穿着一件上面有一个猩红色的“A”的衣服，以表示她犯下了通奸的过错。在有些社会中，小偷被定罪后，手会被砍掉，这无疑是一个代价极其高昂的标签。

背叛者会被打上标记，这种情况甚至也发生在了人类之外的其他物种中。在大海中，医生鱼——裂唇鱼（labroides dimidiatus），既可以选择清除作为它的“邻居”的其他鱼类身上的寄生虫（合作），也可以吃某种更美味的其他食物（背叛）。如果裂唇鱼合作，那么它的“邻居”就可以少受寄生虫的困扰。其他鱼也可观察到寄生虫有没有减少。于是，裂唇鱼“邻居”的清洁程度就成了一个标签，一个代表声誉的具体形象。[7]

连通性与声誉

> 通过声誉机制维持合作有一个条件：个体必须能够知悉自己邻居的行为偏离合作的可能性。为了评估关于这种行为偏差的信息传播出去的可能性，我们可以应用在向传染病模型中加入网络时学到的三个结果。首先，网络的度越大，关于偏离合作行为的信息传播出去的可能性就越大。其次，度分布的变化，特别是超级传播者的存在，也会增加信息传播出去的可能性。再次，如果一个人所背叛的那个受害者，与这个人的其他邻居没有任何联系，那么这个人的邻居就不会知道这个人背叛了他人。因此，要保证声誉扩散，网络必须具有很高的聚类系数，而聚类系数又是社会资本的一个衡量指标。

规则行为者之间的合作

现在，我们放松关于理性的假设，转而假设博弈参与者只会遵循诸如冷酷触发策略之类的规则行事。我们将利用这个更一般的模型探析合作是否可能出现以及如何出现。在这个模型中，假设一群人重复进行囚徒困境博弈，并假设博弈将以如上所述的概率进行下去。我们将证明，如果博弈继续重复进行下去的概率足够高，那么理性的博弈参与者会在这种情况下合作。

与前面那个模型不同，在这里我们假设博弈参与者直接应用特定的行为规则。有些博弈参与者可能会采用冷酷触发策略，有些博弈参与者可能始终合作，而另一些博弈参与者则可能始终背叛。这些策略的某些变体甚至可以在人类之外的其他物种身上看到。雄鸣鸟（Warbler）可能会采取“爱自己的敌人”策略，它们不会大声唱歌，也不会以牺牲邻居为代价来扩大领地。我们可以将这种行为视为一种合作行为。[8]

为了便于解释，我们假设每个个体都与其他人一起博弈。在完成了自己所有的博弈之后，每个人都要公布自己的“成绩”，也就是在博弈中的平均收益。之所以要使用每一场博弈的平均收益而不是所有博弈的总收益来考量，是因为给定博弈以一定概率继续进行，有些博弈参与者（出于偶然）参加的博弈可能会比其他博弈参与者多一些。在这个模型设定中，策略的效能取决于策略的分布。因此，胜出的策略也可能取决于初始分布。如果合作策略在一开始的时候表现是最好的，那么种群中的合作者数量就可能会增加。

在这个例子中，我们随机地向每个博弈参与者分配如下五种遵循行为规则的策略中的一种：始终合作（C）、始终背叛（D）、冷酷触发（GRIM）、针锋相对（TFT）、欺负好人（TROLL）。冷酷触发策略是一开始选择合作，后面也一直继续合作，直到另一个博弈参与者背叛为止，然后就一直背叛。始终合作和始终背叛这两种策略与名字的含义一样：盲目地选择合作或背叛，无论其他博弈参与者的行为如何。针锋相对（或一报还一报）是指在第一次合作，然后每一次都复制另一个博弈参与者在前一次中的行为，两个人都使用针锋相对策略的博弈参与者将永远合作。欺负好人策略则剥削始终合作的博弈参与者，更具体地说，这种策略是，在前两次选择背叛，如果另一个博弈参与者在这两次都没有背叛过，那么就选择永远背叛；而如果另一个博弈参与者在前两次已经背叛过了，那么就先转而在接下来的两次选择合作，然后一直采用冷酷触发策略。

我们首先根据如图22-1所示的囚徒困境博弈的收益矩阵，将每种行为规则策略面对其他每一种策略时的收益计算出来。我们先计算始终背叛这个策略在面对各种策略时的收益。如果面对的是始终合作这个策略，那么它在每一轮博弈中都可以得到的收益为4。与此对应，在这些博弈中，始终合作所能得到的平均收益则为1。如果始终背叛策略“对阵”的是针锋相对策略或冷酷触发策略，那么它在第一轮博弈中获得的收益为4，之后每一轮博弈中都获得2的收益。

如果我们假设博弈会重复多次，那么所有轮次博弈的平均收益将只能略超过2，我们将它记为2^+。而如果始终背叛策略与欺负好人策略相遇，前两轮博弈双方都背叛，然后欺负好人策略在第三轮和第四轮博弈中合作，但是此后一直背叛。因而始终背叛策略还是可以获得2^+的平均收益；同时欺负好人策略的平均收益则略低于2，我们将它记为2^-。

与此类似，我们可以计算出每一对策略的预期收益。[9]表22-1显示了每一个策略对所有其他策略的收益。

表22-1 行策略对列策略的平均收益

博弈参与者的策略	对方的策略				
	始终合作	始终背叛	针锋相对	冷酷触发	欺负好人
始终合作	3	1	3	3	1
始终背叛	4	2	2^+	2^+	2^+
针锋相对	3	2	3	3	3^-
冷酷触发	3	2	3	3	2^+
欺负好人	4	2	3^-	2	3^-

表22-1显示了相互合作、相互背叛和利用其他策略中的缺陷策略的各种情况。仔细研究这张表可以发现，这五种策略中其实有四种是在与自己合作的，因此我们可以将这些策略视为潜在的合作策略。只有针锋相对这一种策略是所有这四种潜在的合作策略都能合作的策略。因此，如果这四种策略的任何一个组合在人口中占了大部分，那么针锋相对策略就能够表现得非常好，尽管不一定总是最好的。[10]

研究者以人为被试，进行了数千次实验，结果充分揭示了人们所选择的策略的巨大异质性。我们将使用表22-1中的收益来考虑给定不同分布时的结果。基于这些策略的不同组合的收益多样性，最优策略将取决于人口的构成。在主要由总是选择合作策略的人组成的人群中，始终背叛这个策略的表现最佳。如果个人选择采用这个最优策略，或者自然选择的作用发挥得非常快，那么人们可能永远无法合作。如果学习或选择以适中的速度发生，那么博弈参与者也会逐渐远离始终合作策略。然而，一旦人口中只包含了很少的采用始终合作策略的人，那么始终背叛策略的表现将不如冷酷触发策略、欺负好人策略和针锋相对策略。这时，这三种策略中的某一种策略将会在人群中扎下根来。无论是在以人类为被试的实验中，还是在计算机仿真实验中，都会发现这种模式的广

泛存在：一开始背叛策略的表现很好，但是不久之后，合作也能扎下根来。我们可以把这些情况下发生的这种事情，称为合作的涌现和合作的演化。

我们不难想象出这五种策略或任何其他策略集上的某种分布，再计算出该分布的平均收益，然后思考通过学习或自然选择接下来可能会发生什么事情。在本书后面的章节中，我们构建了一些关于学习和（自然）选择的正式模型。在这里，我们只是非正式地提出这样的观点，因为我们的目的只是指出合作是否能够出现、取决于种群中的初始战略分布以及人们如何学习或发展新策略。

合作出现或发展的一个必要条件是，合作带来的收益超过了背叛者能够获得的收益。否则，选择和学习都会导致整个种群趋向背叛。为了简化分析，不妨想象如下这个由采取冷酷触发策略、始终合作策略和针锋相对等合作策略的人，以及采取始终背叛策略的人所组成的种群。然后，我们可以计算出，要想让合作策略平均来说表现得更好，必须具备什么条件。这个计算表明，不断发展的合作比简单地维持合作更加困难，而且合作是无法自我引导的，少数合作者无法促成合作的出现。[11]

合作的维持、合作的出现和合作的不断发展，以及合作的自我引导之间的区别值得再三审视。如果当所有参与者都合作时，合作的表现是最好的，那么合作是自我维持的。合作的维持所对应的情况是，通过冷酷触发策略实现的合作是重复博弈的纳什均衡。如果在种群中配对时，合作策略的平均表现优于那些不合作的策略，合作就能够出现或发展起来。

正如刚才已经指出过的，合作出现的条件要比维持合作的条件更难满足。事实上，数学推理告诉我们，以自我引导的方式让合作出现几乎不可能。如果合作者的比例接近于零，合作者的收益就会低于背叛者。这样说并不意味着合作的自我引导永远不会发生，而只是在这个模型中不会发生。为了实现合作，我们需要一部分人从一开始就是合作的。这种情况有可能发生在那些会反思博弈结构的人身上，但是似乎不太可能发生在蜜蜂和树根“身上”。要想理解这种自我引导怎样才能发生，我们需要一些更加精细的模型，以允许局部学习、进化和群体选择。下面就来讨论这些模型。

合作行动模型

为了研究合作怎样才能实现自我引导，在这里引入一个合作行动模型（cooperative action model）。在这个模型中，人们可以采取合作行动，也可以不采取合作行动。[12]合作行动要求个人承担一定成本，会给他人带来收益。在这个模型中，聚类和群体选择都可以产生合作。

合作行动模型与重复囚徒困境博弈之间存在着一些差异。首先，在合作行动模型中，个人并不是两两配对重复进行博弈并在博弈中使用策略、获得收益的，相反，个人要么是合作者，要么不是合作者。其次，合作行动模型不假设理性行为者，也不假设个体会采用更复杂的规则。再次，这个模型中的个体属于一个交互网络。他们的合作行动只会影响与他们有联系的人，也就是他们的邻居。最后，因为个体有固定的类型，所以他们会对所有邻居都采取相同的行动。例如，一个有五个邻居的合作类型的个体，要付出五次合作的成本，并且为另外五个人带来收益。

合作行动模型

> 一个种群由N个人组成，他们或者是合作者，或者是背叛者，连接于一个网络中。在每一次互动中，合作者都要承担合作成本C，而其他人则可以获得合作收益B。背叛则不会产生任何成本和收益。合作优势比率B/C刻画了合作的潜在收益。

在这个模型中，网络发挥了关键的作用。网络的存在，使合作得以出现，甚至可以实现合作的自我引导。一个主要在内部成员之间进行互动的合作者团体或合作者群组会有很好的表现，能够使合作在种群中扩展开来。在生态系统中，后代通常位于父母附近。如果合作者的后代更有可能成为合作者，那么合作的自我引导将会变得更加容易。

为了证明聚类可以导致合作的自我引导，我们从一个只有一部分已经被“充满”的网络开始。这个网络上的每个节点都是一个人可以“居住的住处”。在生物学背景下，这种“住处”就是生物的可行栖息地。然后，我们让合作者或背叛者“住进”网络的一部分。例如，可以先绘制出一个平均度数为10的随机网络，然后在每个节点上掷骰子，如果掷骰子掷出了“6”，就在那个节点上放一个人进去。如果没有掷出“6”，就将这个节点留空。如果我们已经决定要在一个节点上放一个人，那就再掷一次骰子。如果掷出了“5”，就在那个节点上放一个合作者，否则，就放一个背叛者。这个过程结束后，网络中1/6的节点将会被人占有，而且在这

些被占有的节点中，只有1/6是被合作者占有的。

鉴于这个网络结构，每个人的邻居数量将会有所不同，有些人没有邻居，有些人会有四五个邻居。为了在这个网络上实现合作的增长或消亡，我们通过迭代地填充与被占用节点相邻节点的方法来填充网络的其余部分。假设填充空节点的人的类型将与这个节点的邻居中表现最好的那种类型相同（合作者或叛逃者）。图22-3给出了一个线性网络的两个片段，合作者用黑色线条表示，背叛者用灰色线条表示，黑色虚线则表示空节点。每个片段都在中心处包含了一个空节点，它有两个邻居、一个背叛者和一个合作者。在这图22-3中，合作创造的收益为2，而发生的成本则为1。

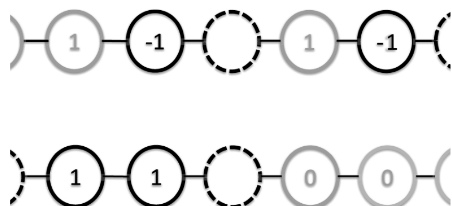


图22-3　两个线性网络中的一个空节点的邻居的收益

在图22-3的上图中，空节点右边的背叛者有一个合作者的邻居，因此可以获得1的收益。空节点左边的合作者有一个背叛者的邻居，因此可以获得-1的收益。根据规则，由于空节点的邻居中，背叛者获得的收益更高，所以这个空节点将由一个背叛者占据。在图22-3的下图中，空节点的背叛者的邻居的邻居也是背叛者，同时，空节点的合作者的邻居则连接到了另一个合作者。在这种情况下，空节点的邻居中以合作者的表现更好，因此空节点将成为合作者。

这个例子表明，一个单独的合作者不能产生一个额外的合作者，但是两个相邻的合作者可以。这就是说，一个小小的合作聚类就可以将合作扩展到空单元上。因此，合作区域可以从少数几个合作者中产生。

我们可以根据相邻的合作者和背叛者的比例以及合作优势比例，写出决定空单元是会成为合作者还是会成为背叛者的更一般的条件。因此，在度数更低的网络中，合作的自我引导更加容易实现。这个发现与我们在分析声誉机制如何维持合作时所得到的结果相反，在那种情况下，更多的连接网络会增加背叛行为破坏某人声誉的可能性，因此更多的连接有助于合作的维持。这也是多模型思维能够产生依赖于特定条件的知识的又一个很好的例子。连通性高的网络能够产生更大的合作还是

更少的合作，这个问题没有单一的答案。如果合作有赖于运用声誉机制的老练的行为者维持，那么连接更多的网络将更有利于合作。如果合作是在不成熟的行为者（如树木或蚂蚁）中自我引导或演化的，那么连接较少的网络应该更能促进合作。

聚类自我引导合作

如果一个空节点的邻居包括了一个合作者（其度数为D且有K个作为合作者的邻居），同时这个空节点的所有非合作者邻居都没有合作者的邻居，那么这个空节点会成为一个合作者，当且仅当合作优势比例高于与合作者数量之间的比例时，即：[13]

$$\frac{B}{C} \geq \frac{D}{K}$$

群体选择

我们要讨论的最后一个自我引导、发展和维持合作的机制是群体选择。这个机制依赖于群体之间的竞争或选择。[14]为了构建群体选择模型，我们将种群进一步划分为若干个群体。在每个群体内，个人的行动满足某种形式的合作行动模型——每个人或者选择合作或者背叛。与以前一样，我们可以认为每个人都有各自的表现。我们还为每个群体分配一个表现，它等于该群体成员的平均表现。群体选择模型假设选择是在群体与群体之间进行的，表现最佳的群体的复制体（副本）将替换表现较低的群体。这种选择有利于合作者组成的群体，它们的表现将会更好。

然而，合作者组成的群体在群体选择时会占优势，这个直觉结论无法回避这样一问题：在任何一个群体内部，背叛者都比合作者有优势。作为例子，不妨考虑两个规模均为10人的群体：第一个群体包含了两名合作者和8名背叛者，第二个群体包含了两名背叛者和8名合作者。如上所述，假设收益等于2且成本等于1。在第一个群体中，每个背叛者的绩效等于4，因为他可以从每名合作者那里获得2的好处；每名合作者的成本为9，并且只获得2的收益，因此其绩效等于-7。第一个群体所有成员的平均绩效等于1.8。在第二个群体中，每个背叛者从8名合作者中的每一个中获得2名，因此其绩效等于16，每名合作者的绩效等于5，因为他从其他7名合作者那里得到14，但是支付9的成本。第二个群体的平均绩

效等于7.2。

这些计算结果揭示了一个矛盾：在每个群体内部，背叛者对合作者有优势，但是表现更好的群体却必须包含更多的合作者。这里的张力是非常明显的：个体选择有利于背叛，但是群体选择却有利于合作。这种张力在各种各样的生态、社会、政治和经济环境下都会出现。例如，让自己的根系与其他树木合作的树木，个体生存条件可能会变得更加糟糕，但是这种合作有助于形成一个更加强大的生态系统，并使之更快速、更有效地扩散到更大的土地上。在一个社区内，合作的个人可以获得的收益少于背叛者，但是合作的社区规模将会扩大。支持自己所属政党的政治家可能比那些只专注于个人支持率的政客更加不容易再次当选，但是凝聚力更强的政党将更有可能发展壮大。在某家企业就职的人如果只专注于学习掌握与本企业有关的知识和技能，对自己可能不利，但是他所属的企业则可能胜过其他企业。

合作行动模型能够帮助我们识别和量化这种张力。为了确定群体选择能否引导、发展或维持合作，还需要往模型中加入更多的细节。为此，特劳森（Traulson）和诺瓦克（Nowak）构建了一个精致的模型。在他们的模型中，种群的人口会增长，而且新出现的成员会复制表现最好的成员的类型。这个模型内置了个体选择和群体选择。选择发生在个体层面，同时表现更好的人更有可能来自合作的群体。当一个群体变得足够大时，它会一分为二，创建出一个新的群体。为了防止种群人口过多，新群体的形成会随机地导致现有的某个群体消失。这最后一个特征引入了一种较弱形式的群体选择。[15]

这些模型证明，群体选择能够增进合作，条件是合作行动的利益相对较大，同时最大群体的规模相对于群体的数量来说比较小。群体选择的效力部分取决于最大群体的规模与群体数量之间的比，这个结果揭示了竞争的必要性。有更多的群体，意味着全部由合作者组成的群体更有可能出现，它也隐含地假设了更多的竞争。最意想不到的一个结果是，最大群体的规模越小，导致的合作更多。较小的最大群体规模可以防止合作者组成的群体被背叛者所支配，也就是说，这限制了个人选择的影响。回想一下在前面举的有8名合作者和两名背叛者的群体的例子中，背叛者的表现更好。如果允许群体的规模扩大为80，那么在群体“分裂”出新的群体之前，它就会包含更大比例的背叛者。如果这个群体在有12名成员时就拆分为两个群体，那么在最坏的情况下，这个群体在拆分时也会包括2/3的合作者。

群体选择拥有增进合作的潜力，这个结论还可以应用到组织内部。大多数组织主要根据个人绩效来分配薪酬。将员工分成若干相互竞争的团队，并根据团队绩效分配奖金和机会，能够诱导合作行为的出现。如果资源流向团队，个人就有动力在这些团队中努力工作，即相互合作。[16]如果合作带来的好处很大，并且团队规模相对于团队数量来说很小，那么这种激励措施应该能够增进团队内部的合作。

但在评估群体选择的潜力时，我们必须仔细考虑个体行动者的复杂情况。树木的适应速度非常缓慢，因此不必要求群体选择很快发挥作用。但是，人类的适应是非常迅速的，因此如果个人背叛的动机很高，那么相应地，群体选择就必须以很高的速度进行。而且，人们也可能会认识到群体选择效果，他们可能会考虑群体之间的竞争，并理解创建一个强大的群体需要符合个体的自身利益。这种认识会使合作更有可能实现。所有这些都表明，我们应该很小心，不要对特定模型中证明的能够促进合作的特定条件过于自信。恰恰相反，我们应该坚持多模型思维，善用判断力，追问定性结论是否仍然成立。

小结

合作如何出现、发展和维持，是来自众多学科领域成千上万的学者一直致力研究的一个难题。模型对这种研究很有帮助，其中最突出的模型是囚徒困境博弈模型。如果我们采用重复博弈框架，并假设行为者是理性的，这个合作难题会暂时消失。模型表明，可以利用进行惩罚的威胁来维持合作。惩罚可以通过重复博弈机制直接实施，也可以通过声誉机制间接实施。这些机制或许可以解释，在利害关系很大且博弈参与者是成熟老练的个体时，合作是怎样产生的；但是它们无法解释为什么蚂蚁、蜜蜂、树木和裸鼹鼠也会合作，而且合作程度是如此之高。当我们考虑了遵循规则的博弈参与者之间的合作时，我们发现要让合作不断发展并不那么容易。理性行为者可以在遵循规则的行为者无法发展合作的环境中维持合作。

我们还发现，像针锋相对这样的简单规则虽然不是最优的，但是可以实现相互合作而且不会被“剥削”。后续研究还表明，如果博弈中会出现随机误差，那么针锋相对策略的表现就不是很好。如果允许出现误差，那么每个使用针锋相对策略的博弈参与者都会在另一个博弈参与者出现误差之后产生一个背叛行动和合作行动组成的循环。如果两名博弈

参与者都因误差意外采取了背叛行动，那么针锋相对策略将导致双方一直背叛，直到另一个误差出现为止。

在现实世界的囚徒困境博弈中，这种误差确实会发生。1983年9月1日，韩国航空公司007航班从阿拉斯加安克雷奇起飞，前往韩国首尔，途中偏离航线进入了苏联领空。一架苏联SU-15战斗机击落了这架民航客机，机上269人全部不幸遇难。美国人认为这是苏联人的“背叛”行径，而苏联人则认为这架飞机在执行间谍任务，认为这是美国人的“背叛”行径。

其他的策略，例如“赢则坚持，输则改之”（Win Stay，Lose Shift）策略，在这样的情况下可以做得更好。在“赢则坚持，输则改之”策略下，相互合作时的收益和诱惑收益都编码为“赢”，另外两种收益则编码“失”。“赢则坚持，输则改之”策略从合作开始，此后，如果赢了，就坚持上一轮所做的一切；如果输了，就转变为另一种行动。只要考察一些例子，你就可以观察到“赢则坚持，输则改之”策略会回归到合作行为上。[17]

我们还描述了另外两种机制。聚类使合作能够实现自我引导。这种机制依赖于合作者之间的互动并通过选择来发展合作。群体选择发挥作用的原理也类似。合作者组成的群体表现得更好并取代了背叛者组成的群体。在正式的模型中，我们发现通过聚类和群体选择实现合作的条件要比通过重复博弈或声誉机制维持的合作更加严格。

我们还了解到，这些机制的成功与否，取决于我们如何对个人进行建模。我们不应期望这些机制对人类、蚂蚁和树木发挥作用的方式完全相同。更精明老练的行为者可能因为他们拥有前瞻能力而能够更好地维持合作，但是，当周围都是合作者时，他们也更有可能发现背叛的好处。

我们的大多数讨论都假设合作是有益的。但是，有些组织也可能通过合作来剥削他人。企业组成卡特尔（cartel）(9)，人为地压低价格；各产油国结成联盟，限制石油产量，以扩张自己的利益，而不管人类的利益是否受到了损害。癌细胞会合作抵抗我们的免疫系统。[18]因此，当我们研究合作的时候，还应该记住，合作不一定是为了共同利益，例如野生水牛并不会受益于狮子之间的合作。

23　与集体行动有关的问题

自从大约5万年前智人发展出了现代创造力、高效和狩猎技能以来，如何可持续地管理环境资源一直是一个非常困难的问题。

贾雷德·戴蒙德（Jared Diamond）

在本章中，我们讨论集体行动问题，也就是那些自身利益与集体利益出现了不一致的情况。集体行动问题在各种情况下都会出现。在机场，乘客可能通过尽可能地挤到行李传送带旁边而获益，因为可以更快地拿到自己的行李，但是如果所有乘客都站到离行李传送带几步以外的地方，那么每一个人都会感觉更好。在美国，人们几乎没有任何动力去成为一个了解有关信息的选民，因为仅凭自己的一票改变选举结果的可能性非常低，尽管如果选民了解有关议题的话，国家的整体绩效将会更好。我们可以把集体行动问题视为一个多人囚徒困境博弈：每个人都有动机去背叛，但是从集体的角度来说，每个人都能够通过合作使自己的境遇得到改善。

学者们经常在历史案例的背景下研究集体行动，例如苏格兰公地的管理或纽芬兰和缅因州沿海龙虾栖息地的演变。[1]其中最著名的案例是贾雷德·戴蒙德讲述的复活节岛上的波利尼西亚人的衰落。[2]复活节岛位于智利以西3 000多千米的南太平洋，方圆1 000多千米内没有其他可以供人居住的岛屿。这个特殊的地理位置决定了，复活节岛上的居民必须自己管理自己。据估计，到17世纪初，复活节岛上的人口已经超过了1.5万人。从16世纪开始，复活节岛上的居民积累了足够多的资源，调集了富余劳动力去建造一种名为莫埃（maoi）的巨型石头人像，它们往往重达80吨以上。

但是，复活节岛上的居民在忙于建造莫埃的同时，却没能共同合作管理好他们的森林。到了1722年，当欧洲探险家第一次登上复活节岛时，岛上的食物已经相当匮乏了，人口也减少到了2 000人左右。岛上已经基本看不到超过3米高的树木了，许多鸟类和动物都已经灭绝。用戴蒙德的话来说，岛上的文明已经崩溃了。当欧洲人带来的病毒害死了岛上几乎所有剩余的原住民时，这个崩溃过程就最终完结了。

根据戴蒙德的解释，复活节岛上文明的崩溃，以及中美洲的玛雅人、美国西南部的阿纳萨齐人（Anasazi）和格陵兰岛的温兰人

（Vinlander）的没落和崩溃，都是由于自然资源的过度开发和气候变化共同造成的，而资源的耗竭，又是因为制度和文化的失败所致。温兰人不但在极地边缘地带放养动物，而且还从原本就非常脆弱的草地上揭起草皮来建造房屋。结果在很短的时期内，他们的土地就因过度使用而变得极度贫瘠了，于是温兰人开始挨饿。与复活节岛上的居民一样，温兰人也未能管理好公共资源。他们砍掉了太多的树木、毁坏了太多的草皮，文明崩溃了。

虽然这些例子都很引人注目，也促人深思，但是由于它们都发生在“遥远”的历史时期，所以许多人可能据此认为，集体行动问题只是在久远的过去才是重要的问题。这是一种相当不幸的框架效应。事实上，由于世界已经变得更加复杂，各个部分之间的相互联系也更加紧密了，集体行动问题其实比以往更加重要。在人类社会的所有方面几乎都面临着集体行动问题。公共教育、身体和精神的健康保障、基础设施、公共安全、司法系统和国防支出，核心都是集体行动问题；管理全球渔业、应对气候变化，特别是减少对大气层的碳排放量，与集体行动问题就更加息息相关了。

此外，由于几乎所有工作都变得更加依赖于团队合作，也必然会产生集体行动问题。员工有很强的动机去“搭便车”——自己偷懒卸责，让别人去努力工作，他们也会对共享工作空间提出过度需求，以确保自己的团队有足够的工作空间。

本章的结构如下。首先定义一个一般的集体行动问题，然后具体分析三类重要的集体行动问题。我们的讨论是从公共物品的供应问题入手的，在这类问题中，个体要为道路、学校和社会服务提供资金，或者为清理公园或水源而付出时间和精力。然后，我们研究了拥挤问题，即个人对诸如道路系统、海滩或公园之类的公共资源的使用必须受到限制。最后，我们分析了可再生资源的开采利用问题——个人需要消费这些可以再生的资源，例如鱼，龙虾和树木。拥挤问题每天都会发生。如果太多的汽车塞满了城市的街道，那么该市或许可以通过向进入城市的汽车征收拥堵费来解决这个问题，因此以往的过度使用不会产生长期影响。然而，森林被过度采伐了、渔场被过度捕捞了，却至少需要数十年时间才能恢复元气。这就是说，我们不得不承担过去合作失败的恶果。

个体激励与集体目标之间不一致的性质各不相同，因此问题的解决方案也不同。我们可以通过征税来解决公共产品供应问题，在某些情况下，也可以通过分类来解决这个问题。拥挤问题则可以通过收费或实施

使用限制来解决，解决可再生资源问题则需要监督、制裁，还需要建立健全解决冲突的机制。

说到底，我们在这里给出的解决方案只是一些最基本的见解，它们必须因地制宜地应用于各种具体情况。任何真实情况都包括了模型无法完全覆盖的复杂层面。巴厘岛的水神庙能够解决水资源分配问题。水资源的分配是一个序列拥挤问题：上游的人可以先利用水资源，然后是中游的人，最后才是下游的人。国际捕鱼权制度通过限制进入的方法解决了可再生资源的公地悲剧问题，不然的话，挪威限制本国渔民近海捕捞的努力，可能因瑞典、俄罗斯和丹麦等国渔民在近海的过度捕捞而无效。[3]在现实世界中，任何一个解决方案的有效性，都部分依赖于我们在第22章中讨论过的关于如何在囚徒困境博弈中实现合作的机制：重复行动、声誉、网络结构和群体选择等。群体选择是间接发挥作用的：成功解决了这些问题的社区和国家将繁荣发展，它们的成功会被其他人所复制。

集体行动问题

在集体行动问题中，每个人都可以在做贡献与免费搭便车之间进行选择。搭便车符合个人利益最大化动机，因为这能为个人带来更高的收益。然而，当每个人都做出贡献时，整个群体能够获得更大的收益。

集体行动问题

在集体行动问题中，N个人中的每个人都要选择是搭便车（f）还是为集体行动做贡献（c）。个人的收益取决于自己的行动和合作者的总数。个人可以通过搭便车获得更高的收益，即，收益（f，C）$>$收益（c，$C+1$），但是当每个人都做出贡献时，所有人的收益总和实现最大化。

我们可以把集体行动问题视为一个多人囚徒困境博弈。因此，我们可以参考第22章中提出的解决方案，来解决如何形成合作和维持合作的问题。然而，那些方法是不完整的，原因如下：第一，集体行动问题涉及群体和社区，而不仅仅涉及配对博弈的个体；第二，许多集体行动问题都有特定的形式，因此其解决方案必须量身定制。

集体行动问题一：公共物品供应问题

我们要讨论的第一个具体的集体行动问题与公共物品的供应有关。公共物品满足非竞争性（non-rivalry），也就是一个人对公共物品的使用不会影响任何其他人的使用，以及非排他性（non-excludability），也就是无法禁止个人使用公共物品。公共物品包括清洁的空气、国防、龙卷风到来之前的警报信号和知识。美国宪法列举了政府有建立司法体系、确保国内安全、防御外国侵略的责任，这些也属于公共产品。

私人物品，比如自行车、燕麦饼干和量角器，既不是非竞争性的，也不是非排他性的。但知识既是非竞争性的，也是非排他性的。比较一下燕麦饼干和三角函数知识，就可以将这种差异突显出来。老师可能会说："梅丽莎，我很抱歉，卡拉已经把最后一块燕麦饼干吃掉啦。"但他永远不可能说："梅丽莎，我很抱歉，卡拉已经用过了毕达哥拉斯定理，现在它永远消失啦。"

公共物品的非排他性和非竞争性导致了集体行动问题。这个问题的出现，不是因为人们不想做出贡献，而是在于人们低估了贡献的价值。每个人贡献出来的每一美元，都可以增加每个人的效用。我们在这里将给出一个正式的模型：每个人都要将自己的收入在一种公共物品和一种代表性的私人物品之间进行分配，可以把私人物品想象成可以花在任何其他东西上的钱。我们可以将这个模型扩展到为包括多种公共物品和多种私人物品的情形，但那只会使分析复杂化，并没有特别大的意义。

公共物品供应问题

有N个人，每人要将自己的收入I（$I>N$）配置到一种公共物品（PUBLIC）和一种私人物品（PRIVATE）上，每个单位的成本为1美元。每个人都有以下形式的效用函数：

$$\text{效用}(\text{PUBLIC},\ \text{PRIVATE})=2\sqrt{\text{PUBLIC}}+\text{PRIVATE}$$

社会最优配置：PUBLIC$=N$（如果$N=100$，那么每个人捐献100美元）。

均衡配置：PUBLIC$=1/N$（如果$N=100$，每个人捐献0.01美元）。[4]

在这里，我们把效用模型描述为公共物品的凹函数和私人物品的线

性函数。对于这样的假设，我们必须说明其理由。凹性对应于收益递减：一个人对某种东西估价随着对它消费的增多而减少。效用对公共物品数量的凹性，意味着公共物品的边际收益递减。这是一个标准假设。人们因高速公路新增的第三条车道而获得的好处，大于因新增第四条车道而获得的好处。污染严重的空气变得清洁对人们的好处要比将清洁的空气中最后一点微尘彻底清理干净更大。

之所以假设私人物品的效用是线性的，是因为它其实代表着所有私人物品的组合。虽然实用函数对任何一种私人物品可能都是凹性的，无论是巧克力、电视机还是牛仔夹克，但是它对于所有商品来说则更可能接近于线性。这个假设还有一个额外的优势，那就是，会使模型更加容易分析。

先求解社会最优配置问题。我们把社会最优配置定义为能够最大化整个种群效用总和的配制，也就是大多数人的最大幸福。[5]社会最优配置要求，每个人都得为种群中的每个人捐献一美元用于提供公共物品。这里需要注意的是，每个人对公共物品的贡献随着种群规模的增大而增加。这个结果不依赖于我们选择的效用函数的形式，因为种群规模越大，能够因非竞争性的公共物品而受益的人越多。这也就是说，享受清洁空气或国防安全的人越多，应该提供的公共物品就越多。

均衡贡献等于1除以种群总人数。随着人口的增多，人们有更大的动机去搭其他人贡献的便车。为了说明个中缘由，我们可以看看让人口规模增大1（加入一个人）时会发生什么。这个“新”人从公共物品中获得的效用，与之前的其他人一样。如果其他人的贡献保持不变，那么这个“新”人为公共物品捐献的动力就会比其他人在以往贡献时更弱。因此，他的贡献会比他们更少。另一方面，当他真的有所贡献时，他将使公共物品的数量有所增加，从而又增强了其他人贡献更少的动机。

因此，这个模型表明，随着人口规模的扩大，公共物品供给（不足）问题也会加剧。公共物品的最优水平提高了，但是人们贡献的动力却下降了。从模型中推导出来的最优贡献水平和均衡贡献水平确实依赖于对效用函数形式的假设，但是公共物品供给不足的现象确实是非常普遍的。

这种分析假设是人都是自私自利的，这也是经济模型中的常见假设。不过，来自问卷调查、实验，以及日常观察的证据表明，人们还有涉他偏好（other-regarding preferences），会考虑他人的利益。人们希望他人也能够与自己一样，上好的学校、利用完善的道路网络。我们可

以在模型中加入一个利他主义参数来模拟这种行为，该参数取零值时就对应于经济学家所假设的自利理性行为者，取值1则对应于每个人为他人考虑的程度与为自己考虑的程度完全一样。纯粹的利他主义者，也就是对每个人都一样关心的人，做出的贡献达到了社会最优水平。只要达不到纯粹的利他主义，就会导致供应不足。

计算表明，在规模很大的种群中，人们的贡献水平与最优配置水平之间的比例，近似地等于利他主义参数的平方。虽然公共物品供给不足的程度取决于效用函数的形式，但是这个例子说明了利他主义的局限性。关心他人的程度相当于关心自己的程度一半的那些人，贡献水平只相当于最优水平的1/4。关心他人的程度相当于关心自己的程度一半的那些人，贡献水平则只占最优水平的1/9。

利他主义者提供的公共物品

N人有利他主义偏好，其对总效用的权重为α：

$$(1-\alpha)\times \text{Utility}_j(\text{PUBLIC, PRIVATE})+$$
$$\alpha\times\sum_{i=1}^{N}\text{Utility}_i(\text{PUBLIC, PRIVATE})$$

均衡纯粹的利他主义者（$\alpha=1$）：PUBLIC $=N$

均衡一般解：[6] $\text{PUBLIC}=\dfrac{[(1-\alpha)+\alpha N]^2}{N}$

实例：$\alpha=1/2$：PUBLIC $\approx N/4$

考虑到我们并不生活在一个由纯粹的利他主义者组成的世界中，我们必须寻找其他机制来解决公共物品供给不足的问题，例如税收。政府征税以支付国防、公共道路、教育和刑事司法体系的支出，并提供其他公共物品。不过，确定税额需要包括个人收入和偏好异质性的更精细的模型。人们可以就税额和税率进行投票。空间投票模型预测税率等于中间选民所偏好的公共物品的水平。但是，如果人们有异质性的收入和偏好，这个水平就可能不是社会最优的。

许多公共物品，如学校、道路网络和资源回收系统，可以认为是属于“本地”的。本地社区可以将本社区之外的人排除在这些公共物品的使用者之外；但是在社区内部，这些公共物品仍然是非竞争性的和非排他性的。对于这种本地公共产品，人们可以根据自己的偏好将它们归入不

同的社区。这就是所谓的蒂布特模型（Tiebout model），它为公共物品的供应问题提供了一个可能的解决方案。想要更好的学校、公园、游泳池和公共安全的人，可以投票支持征收更高的税收，以支付提供这些公共物品的费用。而那些不想要这些本地公共产品的人，则可以住进另一个社区，并只需要支付较低的税。

当然，蒂布特模型也不是万能的，它也不可避免地带来了一些成本，其中之一就是社会凝聚力会因此下降。此外，如果高收入人群以这种方式将自己与社会其他群体隔离开来，就会导致较贫困社区的公共物品供应进一步下降，并减少可以传递信息和知识的网络互动。[7]

集体行动问题二：拥塞模型

第二个集体行动问题是拥塞模型，它刻画了道路、海滩、供排水系统等公共物品对个人的价值随着用户数量的增加而减少的现象。任何一个受过交通堵塞之苦的人，都明白这里所说的含义。道路通畅比严重堵塞能够带给我们更多的乐趣和效用。有人估计，美国每年的交通延误成本已经达到了1　000亿美元以上，尤其是洛杉矶和华盛顿特区，人们每年堵在路上的时间超过了60小时。

拥塞模型假设，资源的总量是固定不变的。每一天，每个人都可以选择使用这种资源或不使用。使用资源的个人收益随着其他用户数量的增加而线性地减少。[8]模型中的斜率，即拥塞参数，描述了这种拥塞效应的严重程度。

拥塞模型

在N人中，M人选择使用一种资源。其效用可以写成如下形式：

$$\text{效用}(M)=B-\theta\times M$$

B表示最大收益，θ是拥塞参数。其余（$N-M$）人不使用这种资源，并且效用为零。[9]

社会最优解：$M=\frac{B}{2\theta}$，且效用$\left(\frac{B}{2\theta}\right)=\frac{B}{2\theta}$

纳什均衡解：$M=\frac{B}{\theta}$，且效用$\left(\frac{B}{\theta}\right)=0$

在社会最优解中，使用这种资源的人的数量等于最大可能收益除以拥塞参数的两倍。这些发现与我们的直觉相符，使用一种资源的人的数量应该随着最大可能收益的提高而增加、随着拥塞效应的加剧而减少。在纳什均衡解中，使用这种资源的人的数量恰好是社会最优解下使用这种资源的人的数量的两倍。这时，拥塞变得如此严重，以至于没有人能够获得任何收益。这个结果依赖于不使用这种资源的效用为零的假设。这个发现有一个违反直觉的含义：建一个美丽的公园可能并不能为社区居民带来太大的效用。在均衡状态下，公园将会变得非常拥挤，以至于在公园休憩并不会比待在家里更好。

当一个模型产生的结果与常识相反时，就需要对结果细加思量。人们肯定更希望拥有一个公园，因此这个模型必定是错的。首先，它之所以出错，是因为我们假设所有人都有相同的偏好。如果不同的人对公园的“享受”程度各不相同，那么在有些人无法因公园而获益的情况下，另一些人仍然可能会获得正效用。其次，这个模型假设公园总是会拥挤不堪，但那不可能是事实。再次，人们的替代选择可能是去海滩，而不是待在家里。因此，新公园的出现，可能会使海滩变得不那么拥挤。最后，人们能够从各种体验中获得效用。如果一个城市里有滑板公园、宠物公园和水上公园，那么人们可以在一定时间内从多种不同的体验中获益。

尽管存在这样一些问题，但是这个简单的拥塞模型的主要结果仍然具有一定的说服力。在繁忙时期，拥塞确实会上升到公园所产生的好处不会比任何其他活动更多的程度。拥塞是不可避免的，尽管不会像只有一个公园时那么严重。更何况，即便建成多个公园也不能保证人们能够在多个公园之间实现最优配置。在下面的例子中，有太多的人都挤到了一个更大的那个公园中。

除了创建更多的公园，社区还可以尝试其他解决方案，例如采取配给制或轮换进入制、进行抽签、收取入场费和扩大容量等。配给制的核心是给每个人或每个家庭分配一定数量的资源。这种解决方案适用于像水这样可分割的资源，但不太适用于公共道路。轮换是按时间分割资源的使用权。例如，为了减少空气污染，城市可以限定在某些日子里只允许车牌号为偶数或奇数的汽车上路行驶。但是很多资源，如受欢迎的公立学校，是不能配给或轮换的，在这种情况下，则可以通过抽签来配置资源。

存在多个拥塞性公共物品的情况

M个人去公园1，（N-M）个人去公园2。相对而言，公园2更大更好，为了体现这一点，假设如下形式的效用函数：[10]

公园1：效用（M）＝N-M

公园2：效用（N-M）＝3N-3×（N-M）

社会最优解：M＝N/2，创造出的总效用为N^2；

纳什均衡解：M＝N/4，创造出的总效为$\frac{9}{16}N^2$

对于道路系统，收取通行费是一种普遍的解决方案。伦敦市会收取车辆进入中心城区的费用，世界各地的收费公路也是如此。收费其实是将资源配置给那些愿意为使用资源付出更多钱的那些人，但是这些人也许不是能够从资源中得到最高效用的人。新加坡同时使用了收费和限制进入的方法。每一年，新加坡都会拍卖固定数量的机动车许可证，许可证的有效期为10年，一份许可证的售价往往超过了一辆普通汽车的价格。为了减少高峰期的交通堵塞，新加坡和伦敦一样，还向进入中央商务区的车辆收费。在同等规模的城市中，新加坡的交通状况算得上相当不错，而且新加坡政府还通过上述方法筹集了大量资金，可以用它们进一步发展公共交通。

扩大道路通行能力，能不能结束拥堵？并不一定。当一个城市通过增加高速公路的车道来缓解交通问题时，会使高速公路附近的房子升值，从而形成一种正反馈回路：房子的数量会随之增加，从而使交通量进一步上升，进而需要更宽的道路。这种正反馈回路，与第18章所描述的系统动力学模型中的正反馈回路类似。

集体行动问题三：可再生资源开采模型

第三个集体行动问题是可再生资源开采模型。在这种模型中，个人所利用的资源是能够再生的。这种模型适用于森林、河流、草原和渔业资源。在这些情况下，未来可用资源的数量取决于现在使用的数量。如果使用得太多，资源就可能无法足够快地再生。资源可能无法快速再生这个事实使可再生资源利用问题比公共物品问题或拥塞问题更加脆弱。如果一个城市在某一年内资金不足，无法提供足够的公共照明，那么它还可以在下一年想办法加以改进，并且不会因为这种错误产生长期不良

影响。但是，如果人们过度捕捞或过度砍伐森林，就会付出持久的代价，因为要生产鱼就必须先有鱼，但并不需要用路灯来制作路灯。此外，可再生资源可能是必需品：食物、饮用水和保暖用的燃料。人们需要开采可再生资源才能生存下去。

可再生资源开采模型

令$R(t)$表示第t期开始时的可再生资源数量，再令$C(t)$表示第t期内耗用的资源的总量，g表示资源的增长率。那么，第$t+1$期的资源数量量由以下差分方程给出：[11]

$$R(t)=(1+g)[R(t)-C(t)]$$

均衡消费水平：$C^{*}=\frac{g}{(1+g)}R$

可再生资源开采模型表明，资源的耗用水平有一个临界点。任何高于均衡开采率的资源开采率都会导致崩溃，这一点在上面的正式模型中可以看得很清楚。我们可以将资源总量视为一个“饼”，开采资源相当于从这个饼中咬掉了一口。由于资源是可再生的，它会产生与剩余资源数量成比例的一定资源。如果资源开采水平较低，资源将会增加；如果开采水平过高，就无法通过再生来弥补了。在这两者之间，存在着一个均衡：开采的资源与再生的资源恰恰相等。

如果开采水平超过了均衡水平，那么模型预测，资源总量将加速下降，最终导致非常突然的崩溃。缓慢的下降之后是急剧的下降，这是对那些管理难以准确测量的资源（例如鱼类资源）的人发出的一个严重警告。年度渔获量提供了关于鱼类资源总量的一丝线索，但是它们并不准确。因此，我们不应该感到惊讶：正如贾雷德·戴蒙德在他的关于社会崩溃的书中所描述的那样，北大西洋的鳕鱼捕捞业在现代也出现了崩溃，就像温兰人在历史上所遭遇过的一样。鳕鱼捕捞业在北大西洋已经有500多年的历史了。最早来到加拿大海岸的英国探险家们讲述了一些关于鳕鱼的“神话”：把一个篮子抛进海，就可以提上一篮子鳕鱼上来；浅滩上聚集了大量鳕鱼，连小船都划不过去！然而到了1992年，加拿大政府不得不严令暂停鳕鱼捕捞。[12]

可再生资源开采模型假设，它们会以不变的速度增加。有了这种假设，我们就可以求出均衡开采水平。但是在现实世界中，可再生资源的增长率每年都在变化。对于牧场来说，增长率取决于温度和降水量；对于鱼类种群来说，增长率取决于可用食物的数量，而食物的数量又取决

于气温变化或气候变化。

在另外两个模型中，这种变化不会产生长期后果。在有些年份，我们可能有比较充裕的公共物品，拥塞情况不会太严重；而在有些年份，公共物品供给不足，拥塞情况就会非常严重。这些变化确实会影响效用，但是影响不会比天气等不可避免的变化更长。但是在可再生资源开采的问题上，情况就完全不同了。如果行为不变，细微的变化就可能导致崩溃与丰裕之间的天壤之别。在图23-1中，我们假设平均再生率为25%、资源的数量为100个单位。给定这些假设，均衡开采水平等于每年20个单位。这幅图给出了介于20%至30%之间的随机抽取的多个增长率下的长期结果。这个模型还假定，资源的最高水平为150个单位。

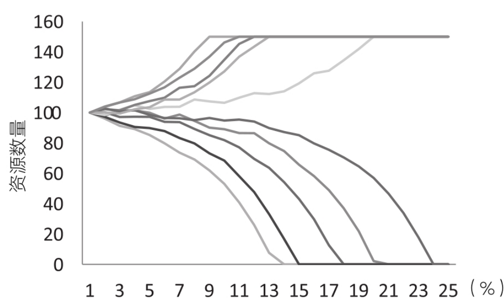


图23-1　资源增长率可变的10种可能路径

在大约一半的路径中，资源水平崩溃了。在另一半路径中，资源的水平增大到最大可能水平。这些变化是不能相互抵消的，恰恰相反，增长率下降或提高的影响会随着时间的推移而累积。[13]从这个模拟实验的结果中，我们可以看出，最优开采政策要求，当某一年出现了资源增长乏力的情况后，下一年必须减少开采，以防止崩溃。

由于可再生资源增长率的变化要求开采率跟随资源增长情况而变化，我们知道管理可再生资源必须做到随时调整开采水平。用来实现这种调整的方法或机制取决于资源的特征。俗话说得好，“世上没有灵丹妙药”。[14]不存在适用于任何情况下的解决方案。到底怎么解决这类问题，取决于资源和管理当地本身的特征。

例如，鱼与牛就不同。管理着一块共享公共草地的多个牛群的社群，可以监督每个牛群的放牧者的行为和资源的水平，即草的数量。过度放牧问题可以通过分配给每个牛群一定时间或一小块草地，或通过轮换放牧等方法来解决。而且，配给或轮换的具体形式，都可以根据牧草的情况随时加以调整。但是，对于一个捕鱼的社群来说，管理资源就需

要更精细的机制来精确地监控个人的行为，因为海中鱼的数量是不可能准确计算的，只能根据捕获量来估算。因此渔业资源开采问题比养牛的不确定性程度更高。管理共同的水生资源，需要更严格的保护和更严密的监测。

集体行动问题的发生

在集体行动问题中，自利行为导致的结果与个人的目标是不一致的。如上所述，这类问题在各种各样的情况下都会发生：在为使用非竞争性和非排他性的公共物品付费时，在决定什么时候将车子开上高速公路时，等等。甚至在我们决定如何在高速公路上驾驶汽车时，也会出现这类问题。在繁忙的高速公路上，如果发生了事故，有些驾驶员会自顾自地打电话、设路障，而全然不考虑他们这种行为对后面所有车辆造成的困扰。

这类问题也会发生在各个层面上。它们会出现在家庭内部：打扫房屋、做晚餐、购物，以及为外出度假存钱，所有这些都可能使个人动机与集体福利不一致。它们还会发生在社区、地区和国家等层面上，只要涉及公共产品的供给以及有限资源的使用和管理就会发生。它们还会发生在全球层面上，例如碳排放问题。大多数国家都希望全球排放水平更低，但是自己又想生产更多能源，也就是排放更多的碳，这时个体的理性行动与共同利益不一致。

集体行动问题在自然界也会发生。在森林里，树木会争夺光线和水分。如果某种树木演化出更高的树冠或更深的根系，那么它将增大自身的生存机会，但是会对其他的树种造成伤害。树木不可能“通过立法”去防止树冠长得过高或根系伸展太深，它们无法实现“社会”最优解决方案。[15]

当涉及的个人或参与者所属的群体越小、越同质化、信息越透明（采取行动更容易、系统状态可容易监测）时，集体行动问题往往更容易解决。家庭通常能顺利解决集体行动问题，但是国际组织却发现全球合作极其困难。要减少碳排放，需要大量不同的行动者之间的能力合作，而且它们所使用的监测机制是不精确的。解决这样的问题需要协调和执行机制。

历史告诉我们，过度捕捞或过度放牧，会导致可再生资源的崩溃。我们可以将同样的逻辑应用于今天面临的集体行动问题。美国著名政治

经济学家埃莉诺·奥斯特罗姆（Elinor Ostrom）花了数十年时间研究现实世界中解决集体行动问题的各种机制。她发现，除了严密监督各种偏离行为之外，能够有效解决集体行动问题的那些社群有如下共同特点：能够就某些明确的界限达成一致、同意明确界定的规则、授权实施渐进式制裁、拥有解决纠纷的机制。[16]

24　与机制设计有关的模型

制度的目的在于改变人类行为。为了保证制度不会随着时间流逝而归于无效，制度必须适应制度所要规范的环境或社会的变化。

珍娜·贝德纳

在本章中，我们讨论如何运用模型去设计政治制度和经济制度。任何一个制度都要包括两个因素：一是人们用来交流信息的渠道，二是人们用来根据所揭示的信息做出决策、重新配置资源或安排生产的程序。在市场中，个人和企业通过价格进行沟通以执行交易并做出生产决策。在层级机构中，人们通过书面语言进行交流以组织工作并推进计划。在民主国家，人们通过投票进行沟通。投票规则决定政策。良好的制度可以促进沟通和行动，从而产生理想的结果，效率低下的制度则相反。

在本章中，我们给出了一个通常称为机制设计（mechanism design）的对制度进行建模的框架。这个框架强调了真实制度的如下四个方面：信息，指参与者知道些什么及应该向他们揭示什么；激励，即采取特定行动的利益和成本；集结，个人行为如何转化为集体结果；计算量，这是对参与者认知能力的要求。

机制设计思想起源于分析资源配置的一般问题，即中央计划体制和市场机制究竟哪一个才能实现资源的最优配置。在早期的模型中，建模者会先给出若干行为规则，例如在市场中充当价格接受者，或者如实投票；然后再研究这些行为的含义，也就是看这些微观行为是怎样集结为集体结果的。后来，这种方法被放弃了，取而代之的是假设人们采取最优化行为的方法。这种假设使这些问题很适合运用博弈论的工具来分析。这就是机制设计的思路。在构建了博弈模型后，机制设计专家求解出纳什均衡，然后在共同的行为假设的基础上进行制度比较。

事实证明，这个分析框架是很有用的。它可以用于搜索现有规则和程序中的缺陷、解释制度成功或失败的原因，可以用来预测结果，还可以用来设计各种各样的制度，包括第2章中描述的频谱拍卖，以及在线交易体系、政府投票系统，甚至还可以用来设计为航天飞机上的研究项目分配空间的程序。[1]

本章讨论的内容包括六个部分。在第一部分中，我们首先使用芒特-

赖特尔图（Mount-Reiter diagram）描述了机制设计的框架。在第二部分中，我们研究了三个人在两个备选项之间进行选择的问题。在第三部分中，我们分析了三种拍卖机制，发现所有这些机制都产生相同的结果。在第四部分中，我们证明上述结果不是一个巧合，并描述了一个基本结果，即收入等价定理（revenue equivalence theorem）。收入等价定理的含义是，只要满足某些假设条件，任何拍卖机制都会产生相同的结果。在第五部分中，我们对多数投票平均分担机制（majority-vote equal sharing mechanism）与枢轴机制（pivot mechanism）进行了比较，它们都是决定是否进行某个公共项目建设的方法。最后，我们沿着在对纳什均衡进行批评时引入的思路，进一步扩展了对机制设计问题的分析。

芒特-赖特尔图

一种机制由六个部分组成：一个环境（世界的相关特征）、一个结果集、一个行动集（也称为消息空间），一个行为规则（人们根据这个规则来做出行动）、将行动映射到结果的结果函数，以及将环境映射到一组希望得到的结果的社会选择对应（social choice correspondence）。社会选择对应通常包括能够最大化参与者效用的结果或帕累托有效（pareto efficiency）配置的集合，当且仅当不存在每个人都喜欢的其他结果，某个结果是帕累托有效的。帕累托效率是下限标准。

帕累托有效

在一个结果集中，对于某个结果，如果存在每个人都喜欢的另一个备选方案，就说这个结果是帕累托占优的。相对应，所有其他结果都是帕累托有效的。[2]

芒特-赖特尔图以图形方式描述了一个机制的上述基本组成部分（见图24-1）。需要注意的是，芒特-赖特尔图将我们想要的东西和已有的东西并列在一起。图的顶部是社会选择对应，它描述了我们希望得到的规范的结果。在图的底部，则列出了我们在现实世界中能做的事情——人们应用他们的行为规则来发送消息或采取行动。结果函数将这些行动映射到结果中去。在理想情况下，下部这个更加复杂的路径会产生与上部路径相同的结果，即人们期望得到的结果。

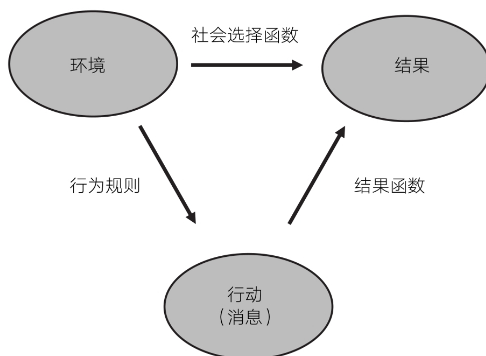


图24-1 芒特-赖特尔图

当然，并不是所有机制都能取得成功。例如，如果环境是由对某种公共物品有偏好的人组成的，那么社会选择对应会将他们的偏好映射到这种公共物品的最优水平上。然而，正如我们在第23章中已经看到的那样，自愿捐献机制，也就是人们按照自己的意愿支付公共物品费用的机制会导致每个人只提供$1/N$单位公共物品而不是最优的N单位公共物品。当一种机制产生的结果与目标不一致时，就说这种机制未能实现社会选择对应。

我们希望机制满足哪些性质？答案必须“因地制宜”。我们在这里只描述其中五个性质。第一，我们会希望这种机制的均衡结果与社会选择对应一致（帕累托有效）。第二，在理想情况下，参与者将会采用占优策略，即他们的最优行动不依赖于他人的行动。如果是这样，就说有效的结果是占优策略可实施的。第三，我们不想强迫人们参与这种机制（自愿参与）。第四，如果这种机制涉及资源的转移或支付，我们不希望增加额外成本或破坏资源（预算平衡）。在本章的后面，当我们分析决定某个公共项目的机制时，会发现这个性质往往很难满足。第五，在许多情况下，我们希望参与者讲真话。我们希望人们发送的消息能够揭示他们的真实信息或真实类型。博弈理论家把这个性质称为激励相容（incentive compatibility）。在大多数有趣的情况下，通常没有任何一个机制可以满足所有这些要求。因此，机制设计理论的一个重要贡献就在于证明什么是可能的、什么是不可能的。

多数规则和拥王者机制

我们要考虑的第一种情况是，人们如何通过投票决定是否采取联合

行动或通过某项法案。假设，乌玛、维拉和威尔三个人想要一起去看电影，在出门前必须决定到底是看动作片、剧情片，还是喜剧片。类似的问题也会出现在军事决策中，比如三位高级军官，要决定到底是主动出击敌军、守卫阵地，还是割让土地。在这两个决策问题中，环境都由三个人组成，其偏好定义在三个备选项上。我们使用排序来表示他们偏好。排序动作片 > 喜剧片 > 剧情片，对应于最喜欢的是动作片，其次是喜剧片，然后是剧情片。我们假设以下的偏好排序：

乌玛：动作片 > 喜剧片 > 剧情片

维拉：喜剧片 > 剧情片 > 动作片

威尔：喜剧片 > 剧情片 > 动作片

在这个例子中，我们将社会选择对应设定为帕累托有效的选择集合。按照上面给出的偏好，喜剧片和动作片是帕累托有效的，而剧情片则被喜剧片占优。

我们首先评估作为一种机制的多数规则。在出现平局的情况下，我们假设选择是随机的。如果人们如实投票，那么喜剧片将会获得两票。然而，假设维拉和威尔都认为另外两个人将会在剧情片与动作片之间出现分歧，而且每个人都会投票给剧情片。再假设投票是连续的，维拉率先投票并选择了剧情片。威尔第二个投票，并同样选择了剧情片。这里，乌玛投什么票其实已经无关紧要了，但是为了避免冲突，她也选择了剧情片。这三张选票构成了一个纳什均衡，也就是说任何一个人都没有动力改变自己的投票。在这种情况下，多数规则并不总是能实现帕累托有效结果。

接下来考虑一下拥王者机制（kingmaker mechanism）。[3]在这个机制中，先在群体中随机地选择一个人当“拥王者”。然后，由“国王”选择一个“国王”来决定这个群体的选择。如果威尔成了“拥王者”，那么他必须在乌玛和维拉之间选出一个“国王”来。他选择的任何人都将成为“国王”并决定大家一起去看什么电影。

如果是一个理性行为者，那个被选为“拥王者”的人会选择自己最喜欢的电影，由此得到的结果将是帕累托有效的。因此，拥王者机制能够实现帕累托有效的结果。这个机制还有一个额外的优点，那就是，如果任何两个人最喜欢的电影是相同的，那么这个机制肯定会选择这个结果。逻辑如下。假设威尔是“拥王者”。如果乌玛和维拉最喜欢的电影类

型相同，那么无论威尔选择谁当“国王”，都会选出这种电影来。另一方面，如果是威尔和乌玛最喜欢的电影类型相同，那么威尔应该选择乌玛当“国王”。

三种拍卖机制

现在我们已经对机制有了最基本的印象了。接下来，我们转向对拍卖的研究。由于像eBay这样的在线拍卖市场的普遍存在，大多数人对拍卖已经有一定的了解。当然，拍卖也广泛用于其他环境，包括政府采购、二手车市场和大多数网络广告。我们在这里将注意力限定在卖家只有一个而竞买人有很多个的情形。拍卖标的可以是房屋、汽车、足球比赛门票或艺术品。再假设每个竞买人赋予拍卖的价值都是唯一的，目的是排除出现平局的可能性。帕累托有效结果是指拍卖标的转移到了对它（们）估价最高的竞买人手中这种结果。任何其他结果，都是帕雷托占优的结果。接下来，我们比较三种类型的拍卖：出价递增拍卖（ascending-bid）、第一价格拍卖和第二价格拍卖。

在出价递增拍卖中，拍卖师喊出一个价格，任何愿意支付这个价格的竞买人都会举手。然后拍卖师不断提高价格，直到只剩下一个竞买人为止。拍卖成交后，最后这个竞买人要支付的是出价第二高的竞买人的价格。在出价递增拍卖中，理性的竞买人会一直参与拍卖，直到价格达到他心目中的价值为止，而他会在那一点上退出拍卖。一方面，在价格达到竞买人的价值之前退出，就有可能无法以该价值的价格买下拍卖标的；另一方面，在价格超过竞买人的价值之后，仍然继续参与拍卖则会带来风险，这意味着竞买人或许可以买下该标的，但是支付的价格却超过了价值，从而导致净损失。

如果所有竞买人都理性行事，那么具有最高价值的竞买人赢得拍卖标的并支付相当于价值第二高的竞买人的价格。这一点可以通过下面的例子来说明。假设有三个竞买人，拍卖标的对他们的价值分别为30美元、60美元和80美元。当拍卖师喊出的价格超过了30美元时，第一个竞买人就会退出。当价格达到了60美元时，第二个竞买人没有理由继续提高出价。因此，第三个竞买人在这个拍卖中胜出，并付出60美元，或者也可能再多一点点。为了简化分析，我们假设第三个竞买人要付出60美元。[4]

在第二价格拍卖中，每个竞买人都以密封的方式提交出价，其他所

有竞买人都不知道这个竞买人的出价金额。拍卖标的由出价最高的竞买人获得，但是竞买人只需支付相当于第二高出价的金额。第二价格拍卖这种机制设计能够保证说实话是最优行动。想象如下这个拍卖的例子。一个对拍卖标的估价80美元的竞买人要决定在第二价格拍卖中如何出价。不妨假设其他竞买人已经提交了出价，那么这个竞买人必须考虑三种可能的情况：其他竞买人的最高出价可能低于80美元、可能等于80美元，也可能超过80美元。不管在哪一种情况下，竞买人如实出价都是最优行动。

通过一个数值实例，上述逻辑会变得更加清晰。我们先假设拍卖标的对这个竞买人的价值为80美元，再考虑其他竞买人提交的如下四种最高出价：70美元（明显低于其价值）、80美元（等于其价值）、85美元（略高于其价值），或90美元（明显高于其价值）。然后进一步假设这个竞买人考虑了四种出价方案，从65美元到95美元不等，如表24-1所示。

表**24-1**　当标的对竞买人的价格为**80**美元时，作为出价函数的净收益

其他竞买人的最高出价	对拍卖标的估价为 80 美元的竞买人的出价			
	65 美元	80 美元	85 美元	95 美元
70 美元（明显低于）	0	+10	−10	+10
80 美元（等于）	0	0	0	0
82 美元（略高于）	0	0	−2	−2
90 美元（明显高于）	0	0	0	−10

从表24-1中可以看出，出价80美元总能带来与任何其他出价一样高的收益。这就是说，按拍卖标的对自己的真正价值如实出价，始终是最优行动（占优策略）。同样的逻辑也适用于所有竞买人，因此所有人都应该按他们的真实价值出价，所以这种机制是激励相容的。因此，在第二价格拍卖中，具有最高价值的竞买人会在拍卖中胜出，并且支付的金额等于价值第二高的竞买人的价值。

在第一价格拍卖中，每个参与拍卖的竞买人都提交一个出价，最高出价者胜出，并且付出的金额就等于该出价。与第二价格拍卖一样，所有参与拍卖的竞买人都同时提交出价，因此没有人知道其他人的出价。参与拍卖的某个竞买人在第一价格拍卖中的最优出价策略，则取决于该

竞买人对于其他竞买人的价值以及可能的出价的信念。我们一般假设，竞买人无法知悉其他竞买人的准确价值，但是他们确实拥有关于这些价值分布的正确信念。我们假设竞买人的价值均匀分布在0美元至100美元之间，并且所有竞买人都知道这种分布。竞买人也知道拍卖标的对他们自己的价值。

通过简单的微积分计算，我们可以证明，如果竞买人的出价是均匀分布的，并且如果所有竞买人都以最优方式出价，那么在只有两个竞买人的情况下，每人的出价都应该是自己真实价值的一半；在有N个竞买人的情况下，每个竞买人的出价应该是自己价值的N-1/N。这也就是说，与其他19个人一起参加拍卖的竞买人应该按自己真实价值的95%出价。根据这个出价规则，具有最高价值的竞买人肯定能够买下拍卖标的。我们还可以证明，这个竞买人付出的金额等于价值第二高的竞买人的预期价值。因此，第一价格拍卖也能够产生帕累托有效的结果，且拍卖成交价格与对拍卖标估值第二高的竞买人的预期价值相对应。[5]

在写下这个模型之前，我们许多人就已经知道，参与拍卖的竞买人越多，任何一个竞买人的出价就应该越高。但是，不利用模型进行推导，我们就不能得出均衡出价规则。这里第一价格拍卖模型为我们提供了一个精确的公式。这个公式告诉我们，出价会随拍卖标的对竞买人价值的增加而提高，而这就意味着具有最高价值的竞买人将赢得拍卖，就像在其他两种拍卖中一样。

收入等价定理

在三种拍卖机制的任何一种中，都是具有最高价值的竞买人在拍卖中获胜。因此，所有这三种拍卖机制都产生了有效的结果。此外，中标的竞买人付出的预期金额等于第二高出价者的价值。换句话说，所有这三种拍卖都导致相同的预期收入并将拍卖标的分配给了同一个竞买人。这一点非常引人注目。而且，更加值得注意的是，我们还可以证明，在竞买人采取最优行动、且最高出价者在拍卖中胜出，同时价值为零的竞买人不会得到任何收益的任何拍卖中，胜出者和预期收入都是相同的。换句话说，所有拍卖机制都会产生相同的预期结果，这个结论就是通常所称的收入等价定理。[6]

收入等价定理

在任何拍卖中，如果竞买人从已知的共同分布中抽取独立私人价值，那么只要满足如下条件，拍卖就必定会给卖方带来同样的收入、给买方产生同样的预期收益：每个竞买人的出价都是为了最大化自己的预期收益、最高出价的竞买人总能赢得拍卖标的，同时价值为零的竞买人的预期收益为零。

收入等价定理还意味着全支付拍卖（all-pay auction），在这种拍卖中，每一个竞买人，即便是未成功的竞买人，都要按自己的出价付款，这也会产生与第二价格拍卖相同的结果。[7]即便像第三价格拍卖这样“几近疯狂”的拍卖设计（最高出价者胜出并支付相当于第三高出价的金额），也会产生相同的胜出者和相同的收入。

当然，收入等价定理并不意味着拍卖规则是无关紧要的。在拍卖中，竞买人可能不一定使用最优策略。或者，在第一次价格拍卖中，他们可能会对其他竞买人的价值分布有不同的信念。只要出现了任何一种情况——竞买人不进行最优化或存在不同的信念，那么各种拍卖的收入就可能会有所不同。经验证据和实验确实表明，理论与现实之间存在着一些差异。

但是，正如我们前面在讨论什么情况下可以预期行为者理性行事时所指出的，利害关系越大，竞买人越精明老练，他们理性行事的可能性就越高。在消费品的在线拍卖中，我们会预期一些人可能会遵循经验法则或者受认知偏差影响做出不理智的行动（例如每次以10美元增价竞拍）。但是，在价值千百万美元的石油租赁拍卖中，竞买人肯定会事先做好准备，了解全部可以掌握的信息，并学会必要的技能。

此外，拍卖形式本身也可能会影响竞买人的数量。在木材拍卖中，第一价格拍卖所能吸引的竞买人数量要比出价递增拍卖更多，因为如果大竞买人（大公司）提交的出价较低，小竞买人也有机会胜出。但是，在出价递增拍卖中，小竞买人是没有机会胜出的，因为大竞买人可以观察到小竞买人的出价并给出更高的出价。[8]

不同的拍卖形式对参与者认知能力的要求也不同。在某些拍卖形式中，最优行为是很容易学习的。例如，在出价递增拍卖中，竞买人应该一直参与拍卖，直到价格达到自己的价值为止。其他未遵循最优策略的竞买人可能会导致采取最优策略的竞买人获得更高或更低的预期收益，但是他们还是不会改变最优策略：只要价格低于自己的价值，竞买人就应该继续参与拍卖。与此类似，在第二价格拍卖中，竞买人应该始终遵

循同样的策略，按自己的真实价值出价。但是，要真正理解按自己的真实价值出价是最优策略，需要多个逻辑步骤。

请回想一下占优策略的定义：无论其他人采取什么策略，占优策略都是最优的。出价递增拍卖和第二价格拍卖中都存在占优策略。但是，第一价格拍卖则不存在占优策略。在第一价格拍卖中，一个竞买人出价策略的变化可以改变另一个竞买人的最优策略。如果一个竞买人的出价总是为0或50，那么另一个竞买人就应该总是出价1或51，后者没有理由出价60或70。考虑到那个竞买人的出价行为，出价60肯定能够赢得拍卖标的，但是51的出价也已经足够了。

而且，即便某个拍卖中存在占优策略，也并不是所有的占优策略都同样容易推断。在出价递增拍卖中，只要价格低于竞买人的价值就一直继续参加拍卖，这个最优策略并不是一目了然的，而是至少需要通过一步推理才能得知的：如果价格低于你的价值，请以这个价格买下。在第二价格拍卖中，竞买人必须考虑过多种可能发生的情况才能看清楚按自己的真实价值出价是最优策略。当然，如果一个人参加过好几次第二价格拍卖，那么他应该能够学会在未来的拍卖中采取这种策略。

关于各种形式的拍卖，要考虑的最后一个问题是，拍卖是不是会鼓励非最优行为。在第一价格拍卖和第二价格拍卖中，竞买人要在不知道其他竞买人出价的情况下提交自己的出价。而在出价递增拍卖中，竞买人可以直接观察到价格一路上行并知道哪些对手仍然在参与拍卖。这可能会导致竞买人赋予“在拍卖中胜出”这个结果某些附加价值，或者甚至赋予“出价压过别人一头”这种行为某些附加价值。主持慈善拍卖的拍卖师试图通过诉诸情感来调动竞买人提高出价，例如展示儿童在用拍卖所得购买的新游乐设备上快乐嬉戏的视频。

不同的策略能否成功，还取决于竞买人的精明程度。我们无法想象，在木材拍卖会中，竞买人会被人引诱报出严重高于自己预测估值的出价。但是，在慈善拍卖会上出价过高却不算非常罕见。有人说，也许竞买人在投标过程中会改变自己对拍卖标的的估值，但这只是一个猜想，我们只需要承认它确实可能会发生。在第一价格拍卖和第二价格拍卖中，竞买人只有一个出价机会，因而无法在拍卖期间通过感情因素来影响他们的决定。

最后，在第一价格拍卖和出价递增拍卖中，价格等于最高的出价。而第二价格拍卖中，价格等于第二高的出价。这就给人留下了这样的一个印象：在第二价格拍卖中，卖方本来应该可以获得更高的价格，并在

一定程度上解释了为什么政府不愿意使用第二价格拍卖。想象这样一个例子，在美国政府组织的油田开采权的拍卖中，政府收到了三个投标，第一个是600万美元，第二个是800万美元，第三个是1　200万美元，如果采用第二价格拍卖，那么第二天报纸上的新闻标题就可能是："政府得到1　200万美元的出价，但是却以800万美元的价格售出了油田。"那无疑是一场公关灾难。事实上，任何了解拍卖理论的人都知道，如果政府进行的是第一价格拍卖或出价递增拍卖，那么最高出价是不可能达到1 200万美元的，而只能达到800万美元。

正如我们在本书中一直反复强调的，正式的模型可以揭示特定结果成立所需的条件。等价定理并不意味着所有拍卖机制都会产生相同的结果。它只说明，在满足如下条件时，所有拍卖都是等价的：竞买人采取最优出价策略、最高出价者赢得拍卖标的，同时对拍卖标的估价为零的竞买人的预期收益为零。卖方可以通过放宽这三个假设中的任何一个来筹集更多资金。卖方很难让人们违背自身利益行事，而且也可能无法从认为拍卖标的完全没有价值的人那里榨取到钱财。

因此唯一剩下的可能性是，不将拍卖标的出售给出价最高者。而要做到这一点，最直接的一个方法根本不出售拍卖标的。如果卖方知道竞买人的价值分布，那么就可以设定一个保留价格或底价，即最低出价。在某些情况下，这可以增加卖方的预期收入。例如，假设卖方确定，某个拍卖标的对三个竞买人的价值分别为5美元、10美元和60美元。在上述三种拍卖形式中的任何一种中，胜出者的出价都是60美元，而且都只需支付10美元。在这种情况下，卖方可以规定，保留价格为60美元，然后再进行一个第一价格拍卖，从而获得更高的收入。

公共项目决策问题

接下来，我们比较一下用来决定启动建设一个公共项目的两种机制，如新学校、新高速公路或新体育场。我们假设，新项目会给每个人带来一定价值，同时也要每个人付出一定成本。

公共项目决策问题

V_1，V_2，...，V_N表示N个人赋予一个公共项目的货币价值，并假设该公共项目的成本为C。那么当且仅当$C < V_1 + V_2 + ... + V_N$时，这个项目才会启动。

我们首先考虑多数投票平均分担机制（majority-vote equal sharing mechanism）。在这个机制中，个人投票决定是否启动某个公共项目。如果多数人投了赞成票，那么项目启动，而且成本由所有人平均分摊。

多数投票平均分担机制

个人投票表示赞成或反对启动某个公共项目。如果多数人投票支持该项目，那么该项目启动，并且每个人都承担C/N的成本。如下面的例子所表明的，这种机制可能会违背效率条件和自愿参与原则。

从空间投票模型中可以看出，项目是否启动取决于中间选民的偏好。在我们现在讨论的这种情况下，中间选民就是指公共项目对他的价值位于中位数的人。根据定义，这个机制满足预算平衡条件和激励相容要求。但是，这个机制不一定能够满足效率条件和自愿参与原则。

假设有三个人的价值分别是0美元、120美元和150美元，而公共项目的成本是300美元。有效的结果是不应启动该项目，因为300美元的费用超过了个人价值的总和。然而，考虑到成本将平分，每个人将投票决定是否以每人100美元的成本进行该项目。因此，这三个人中有两个将投票支持这个项目，而且这个项目将会启动，这是一个低效的结果。此外，价值为0美元的个人获得-100美元的回报，因此这个例子也违反了自愿参与原则。

接下来讨论公共项目决定的第二种机制，枢轴机制。在这种机制下，每个人都提交自己对公共项目的估值，如果估值总和超过了项目成本，就启动该公共项目，否则就不启动。同时，对某个人征税的金额等于项目成本减去所有其他个人估值的总和。如果其他个人的估值已经超过项目成本，这个人就不用支付任何费用。

枢轴机制

个人i对一个成本为C的项目提交自己的估值$\hat{V}_i$。如果所有个人的值的总和超过了成本，那么就启动这个公共项目，即：

$$\hat{V}_\Sigma = \hat{V}_1 + \hat{V}_2 + \cdots + \hat{V}_N \geqslant C$$

如果$C-(\hat{V}_\Sigma - \hat{V}_\text{i}) < 0$，那么个人$i$不用交税；如若不然，个人$i$就要缴纳数额为$C-(\hat{V}_\Sigma - \hat{V}_i)$的税收。这个机制是激励相容的

（$\hat{V}_i=V_i$）、有效率的，而且个人行为也是符合理性。它还实现了占优策略的有效结果。但是，正如下面的例子所表明的，这个机制可能会违背预算平衡条件。

实例：（V_1，V_2，V_3）＝（60，120，150），C＝300。

这个公共项目本应启动，因为300＜60＋120＋150。个体1要缴纳的税收为30，即总成本减去其他人估值之和的结果（300-270）；个体2要缴纳的税款为90；个体3要缴纳的税款为120。由此得到的总税收为240，低于项目的成本。

这个机制满足激励兼容条件，原理与第二价格拍卖类似。现再举例说明。假设一个公共项目的成本为300美元，而且某个人对这个公共项目的价值估计为80美元。有三种情况需要考虑。如果其他人的估值的总和低于220美元，那么这个人没有动机提交超过80美元的估价，因为到时必须由他来支付该金额。如果在另一个极端上，其他人的估值总和超过了300美元，那么这个人什么都不用付出，他可以给出任何估价。但是，如果其他人的估值总和介于220美元到300美元之间，且这个人提交了80美元的估价，那么他要承担的成本将等于300美元减去那个总和的差，并且这个项目将启动（这是一个有效率的结果）。他将不会提交70美元的估值，因为其他人的估值总和可能是225美元，那么他提交的这个低估值将使这个公共项目无法启动。而如果他提交的估值为80美元，那么他所要承担的成本仅为75美元。

由于枢轴机制满足激励兼容性，因此它也满足有效性。只有在估值总和超过了成本的情况下，才能实施公共项目。需要注意的是，因为报告一个人的真实价值是一个占优策略，所以有效结果也是占优策略可实施的结果。此外，由于每个人最多支付项目对自己的价值，这个机制也满足自愿参与原则。但是，这个机制不一定会得到预算平衡的结果，事实上，这个机制只是在极少数情况下才能做到这一点。

对于决定公共项目的决策问题，任何机制都无法满足我们可能想要达到的所有标准。事实上，当用模型证明了这一点之后，就可以省下很多时间，不会再无谓地去尝试一些不可能实现的事情。正如工程师不会浪费时间去建造永动机一样，机制设计专家也不会去尝试为公共项目决策问题寻找一个满足激励相容、个体理性、有效率且预算平衡的机制。事实上根本不存在这样的机制。

枢轴机制已经相当不错了，但是它不能满足预算平衡条件。而且，

这个缺陷无法通过提高人们为项目缴纳的税额来解决，因为那样做会使这个机制不再是激励兼容的或个体理性的。那样的话，个体会有动机去撒谎，有些人可能会被要求为项目贡献出超过其价值的东西。一种可能的解决方法是通过其他途径来增加税收，以便为项目提供资金池。当然，那种途径本身也会产生激励问题，但那不是直接的。

更好的解决方法是同时拥有其他更多的资金来源。例如，一所大学，如果有规模很大的校级基金，同时组成大学的各个学院又各自都有独立的基金，就可以用这种机制来决定是否建立一个新的学生会。大学每个学院的院长都有动机真实地揭示学生会对自己学院的价值，同时大学校长则可以用校级基金来弥补可能资金的不足。由拥有预算权限的分支机构组成的企业也可以考虑这样做。例如，当这样的企业在决定要不要切换到一个基于云的系统时，就可以采用枢轴机制来决策，而且更高的管理层可以解决任何可能出现的缺陷。

小结

作为一个框架，机制设计理论使我们能够依据各种标准对不同的机制进行比较。机制设计能不能产生有效率的结果？人们会说实话吗？人们会自愿参加吗？某个机制是会产生盈余、还是会导致损失？利用机制设计框架，我们还可以推导出可能的结果。当然，一般来说，我们无法在同一个机制下满足所有想要达到的标准。在进行机制设计的时候，建模者摇身一变，成了工程师。我们使用模型来尝试构建可行的解决方案。

随着技术的进步和变化，机制也会发生变化。以谷歌等互联网搜索网站所使用的广告拍卖算法为例。最初，谷歌是按固定价格收费的，每千次点击收取多少费用。随着信息技术的发展，谷歌能够同时进行数百万次拍卖，在这种情况下，固定收费就不再是一种最优机制了。通过引入拍卖方法，谷歌不仅增加了收入，并且更有效地分配了广告空间。谷歌现在使用的是一种广义的第二价格拍卖。每个竞买人都会提交自己对每次点击的出价，目的是通过特定的关键词进行推广，例如，治疗一种因接触石棉而导致的癌症的药物，最高出价者得到第一个广告位，第二高出价者得到第二个广告位，第三高出价者得到第三个广告位……以此类推。这些竞买人付出的价格则通过第二价格拍卖来确定。

假设前四个最高出价分别是每次点击10美元、7美元、6美元和3美

元。那么，第三高出价者获得第三个广告位，同时支付的价格等于第四高出价者的出价，即3美元；第二高出价者支付的价格则等于第三高出价者的出价，即6美元；最高出价者则要付出7美元的价格。[9]在了解到了广告商的估值后，谷歌还可以设定保留价格（底价）并收取更高的费用。但是，如果竞买人也知悉了谷歌的这些计划，这个结果就不一定是有效的。认为自己有可能成为高出价者的竞买人不希望谷歌知道他的估值。同时设定底价也会损害谷歌的声誉，而且，底价也会被视为一种非合作行为，因为谷歌无法证明自己有权对网页上的位置拥有保留价值。除非卖出去了，否则关键字搜索上的顶级广告位对谷歌本身来说几乎没有任何价值。而对于销售古董家具或二手车的人来说，情况却并非如此。这些物品对卖家是有价值的，因此设定一个底价是合理的。而且，谷歌是一个重视声誉的企业，设定底价可能会激怒广告客户。

总而言之，机制设计框架可以帮助我们设计制度，也能够指导我们在不同的制度之间进行选择。有了这个框架，就可以推断出什么是可行的、什么是不可能实现的。也许，我们很难设计出这样一种机制：既能够产生有效率的结果，又能够引导人说真话，同时还满足预算平衡条件。如果确实是这样，我们就不应该浪费时间和精力去设计这样的机制，而应该将更大的时间精力投入到如何实现（例如）效率和平衡预算之间的权衡上。

我们还可以利用机制设计框架来探索一些更宏大的问题，例如，我们在什么情况下应该利用市场、在什么条件下应该投票、在什么时候应该依靠等级体系，在什么环境中应该转而采用自愿的集体行动来分配资源或采取行动。[10]市场、民主、等级制度和集体行动这四个机制中的每一个，都只在某些环境下运行良好，而在另外一些环境下则表现不佳。例如，我们不会用投票来决定人们购买什么商品，也不想让市场去决定谁来当美国总统。

在社会和组织内部，我们都观察到了这些制度形式。以大学为例。大学要面对一个教职市场，同时又要通过民主程序来雇用教师；要通过等级体系分配课程作业，并运用集体原则制定战略、计划。非营利机构、以获取利润为目标的企业，以及政府机构，也都需要混合运用各种不同的制度形式。利用机制设计工具箱，我们可以比较每种制度的运行方式，然后再把各种任务更好地适配各种制度。

25　信号模型

诚实的人不会隐瞒自己的所作所为。

艾米莉·勃朗特（Emily Brontë）

在本章中，我们研究信号模型（signaling model）。这类模型确定了人们发送“昂贵”的信号以揭示信息或类型的条件。一个人可以通过购买昂贵的艺术品表明自己的财力，通过攀登很高的山峰来展现自己的体力，或者通过在社交媒体上发声支持受难者来表达自己的同情心。利用发送信号来揭示自己的身份一直都是人性的一个部分。

早在19世纪，经济学家托尔斯坦·凡勃仑（Thorstein Veblen）就提出了“炫耀性消费”的概念，大大增进了我们对信号的理解。凡勃仑观察到，人们经常选择通过炫耀性消费来表明他们的社会地位，而不仅仅会购买那些能够带来直接享受或效用的商品。如果活到今天，当凡勃仑看到现代人的炫耀性消费行为，他肯定将会心一笑：例如迈巴赫敞篷车，每辆售价将近150万美元；10年陈酿的克丽丝特尔酒（Cristal），每瓶售价超过1 500美元；徕卡相机，每台售价数万美元……

炫耀性消费由来已久，部分原因在于人类很在意别人对自己的看法。这种消费行为之所以经久不衰，还因为消费可以起到信号的作用。[1]我们不能完全看清某个人，所以我们依赖于他们穿的衣服、开的汽车、喝的酒来推断他们的“隐藏属性”。如果我们看到一个人开着昂贵的汽车，那么大体上可以推断出他拥有一定财富。一个人向慈善组织大笔捐款，表明他是一个慷慨大方的人，因为没有自私的人会做出这样的行为。一个人在社交媒体上宣布自己获得了理论生物学博士学位，那是传达关于他的智力水平和所从事职业的信号。几乎所有行动都在一定程度上传递了某种信号。当政客们投票决定是对某个国家宣战、还是实施制裁时，他们就发出了关于自己意识形态立场的信号。某些有长期目标的政客（例如打算日后竞选总统）可能会试图通过投票发出最有利于自己未来政治前途的信号，因而不一定会给最有利的政策投赞成票。

在本章中，我们首先研究离散信号模型。在这种模型中，个体可以选择发送信号或不发送信号；同时，不同的个体发送信号的成本也不相同。要让信号发挥作用，它们就必须是昂贵的（有成本的）或可验证的。这将是本章的一个重点内容。例如，一个雇主打算从新入职员工中

选一些人夏天到西班牙巴塞罗那出一趟“美差”。所有申请的员工都在简历中声称自己会西班牙语。但是，简单地说自己会说西班牙语只是一个没有成本的信号。为此，雇主可以启动一个“西班牙语争章活动”，要获得一枚徽章，需要用西班牙语完成一个小时的演示。对于真的熟练掌握了西班牙语的员工，发出这个信号，也就是用西班牙语完成演示的成本较低。但是对于那些不懂西班牙语的员工来说，准备长达一个小时的西班牙语演示的成本却高得令人望而却步。用信号模型的术语来说，这个徽章就将熟练西班牙语的人与不懂西班牙语的人区分开来了。

接下来，我们简单地介绍一个连续的信号模型。在这个模型中，信号的大小是可变的。一个夏令营的皮划艇队一般只能有一个领划的皮划艇运动员，人们希望在这个位置上的是一个耐力非常好的人，那么，怎样才能挑选出这样一个人来呢？营地主管可能会要求两个提出申请的选手连续划皮划艇10个小时，以便将皮划艇划到尽可能远的距离上。那两个皮划艇运动员中耐力更强的那个选手可以划到另一个实力较弱的选手无法企及的距离上，从而让自己脱颖而出。

离散信号模型和连续信号模型都能为我们提供信号什么时候分离、什么时候不能分离的条件。因此，它们给我们提供了比文字描述更加准确深入的见解。文字描述能够告诉我们，人、动物、政客和政府发出了什么信号和为什么要发出信号，但是却不能向我们明确表征信号发出的时间和信号的强弱。这些模型还可以非常清晰地解释，为什么学生会如此努力地试图证明自己对学院的价值。在本章的结论中，我们讨论了信号模型的理论贡献及其政策含义，还将讨论生态学、人类学下的信号模型，以及商业活动中的信号模型应用。

离散信号模型

我们从离散信号模型开始讨论。在这种模型中，人们要决定是否采取某种行动。你可以买一块昂贵的手表来证明你拥有大量财富，可以通过主修物理学来证明你智力超群，可以通过横渡英吉利海峡来证明自己身体健康。但是你不能半途而废：要么发送信号，要么不发送。这个模型假设，存在两种类型的人，强者和弱者。这两种类型在现实世界中，可以对应于身体健康有资格进入海军陆战队的年轻人和身体孱弱的人，也可以对应于会两种以上外语的员工和只会本国语言的员工，等等。

发送信号的成本取决于个体的类型。这里所说的信号，可能是为有

可能成为海军陆战队员的申请人提供的为期一个月的魔鬼训练计划，也可能是求职者用西班牙语完成的长达一个小时的演示。强健的准海军陆战队员会发现完成训练计划的成本更低。在模型中，我们假设发送信号的每个人可以平等地分享总收益。对于这个假设，可以从两个角度加以解释。在某些情况下，某种资源可能会在发送信号的所有人之间分配。例如，向学校捐赠了1 000美元的每个人（捐赠是慷慨的信号）的名字都会被刻在一面墙上。在另外一些情况下，例如对于准海军陆战队员和求职者，则可以从发送信号的人的集合中随机挑选一些出来作为“中奖者”。

这个模型支持三种不同的结果：混同（pooling），所有人都发送相同的信号；分离（separating），每种类型的人各自发送一个独特的信号；部分混同，其中一些类型区分开来了，其他类型则没有区分开来。

离散信号模型

一个规模为N的种群由S个强者类型的个体和W个弱者类型的个体组成；这两种类型的个体发送信号的成本分别为c和C，且$c<C$。种群中所有发送信号的成员平均分配B的收益（$B>0$）。这个模型有三种可能的结果：

混同$\left(C<\frac{B}{N}\right)$：两种类型的个体都发送信号。

分离$\left(c<\frac{B}{S}\right.$，且$\left.\frac{B}{S+1}<C\right)$：只有“强者”类型的个体发送信号。

部分混同$\left(c<\frac{B}{N}<C<\frac{B}{S+1}\right)$：“强者”类型的所有个体和“弱者”类型的部分个体发送信号。

在这个模型中，我们假设个体在给定其他个体行动的情况下做出最优选择。也就是说，我们将它视为一个博弈，并求解它的均衡。在混同均衡中，每个人都发出信号，如果收益很高并且弱者类型的人发送信号的成本较低，就会存在这种均衡，确切的条件是收益除以人数的商必须超过弱者类型的人的成本。例如，假设一位捐助者捐赠了100万美元设立了一个奖学金，用于奖励某所高中所有100名毕业生。假设有50名学生属于强者类型，只需每周学习两小时就能够从高中毕业，而另外50名学生则属于弱者类型，每周必须花费10小时学习才能完成高中学业。对于强者类型的学生，我们可以将学习成本估计为2 000美元，而对于弱者类型的学生，学习成本则为5 000美元。如果所有100名学生都顺利毕业，那么每人都可以获得10 000美元的奖学金。因此这两种类型的学生

都有很强的动机去学习。

但是，如果假设我们将奖学金总额减少到了20万美元。现在，如果所有100名学生都顺利毕业，那么每人都只能获得2 000美元的奖学金。这样，学习就不再符合弱者类型学生的自身利益了。而对于强者类型的学生来说，现在每人可以得到4 000美元奖学金了，因此学习仍然是有意义的。但是，这个数额仍然不足以诱导弱者类型的学生毕业，哪怕只有一个都不可能。在这种情况下，奖学金的设置导致了分离均衡。

最后，假设奖学金的总额为40万美元。再一次，如果所有100名学生都毕业，那么弱者类型的每个学生的所能得到的奖学金为4 000美元，低于5 000美元。因此，他们不会全都选择学习。但是，如果弱者类型的学生都不学习，那么强者类型的学生每人将会获得8 000美元的奖学金，这个数额对弱者类型的学生来说也很有吸引力。因此在均衡中，最终将会有30名弱者类型的学生与所有50名强者类型的优秀学生一起毕业。结果是总共有80名学生毕业，每人得到5 000美元的奖学金，这也正是弱者类型的学生的学习成本。我们将这种结果称为部分混同，因为有部分弱者类型的学生与强者类型的学生混同在了一起。

部分混同均衡比其他两个均衡更加复杂，因为它需要弱者类型的学生实现彼此之间的某种协调。我们可以假设存在某个过程，弱者类型的学生会与其他人沟通，告诉别人自己计划采取能够确保毕业的行动。或者也可以假设弱者类型的学生的努力恰恰达到了这样一个水平：他们能不能毕业完全是随机的，并且该努力水平会导致30名弱者类型的学生毕业的期望。第二种情况似乎不那么合理。一般而言，我们应该将部分混同均衡解释为一个基准，即如果人们试图最优化，会发生什么。是否能达到部分汇集均衡，可能取决于具体情况，尤其取决于人们是否可以交流各自预期中的行动。

连续信号模型

在离散信号模型的部分混同均衡中，强者类型的人在有些时候可能会觉得沮丧。如果他们能够发出足够强烈的信号，就可以完全与弱者类型的人区分开来，并且获得更高的收益。为了在模型中包含这种可能性，我们可以改变假设并允许强者类型的人自行选择它们要发送的信号大小。这只需要对模型稍做修改即可。为此，我们将离散信号的发送成本重新解释为连续信号的每单位发送成本。此外，我们假设，对于任何

固定数量的信号，强者类型的人每单位成本更低一些。

为了在这个新模型中实现分离均衡，强者类型的人必须愿意选择一个对于弱者类型的人来说成本极高的信号，当然条件是在考虑了收益和成本之后，这个信号仍然是值得发送出去的。通过模型推导，我们发现至少有一些强者类型是可以分离的，但不一定是全部。

令人惊讶的是，随着强者类型的群体规模的增大，信号的量级反而会变小。这种情况之所以会发生，是因为强者类型的人发送信号的好处减少了。成为规模更大的群体的一部分，能够得到的好处反而会更少。完全分离这个条件意味着，当强者类型的人数很少，或者强者类型的人发送信号成本要比弱者类型的人低很多时，分离均衡更有可能实现。

连续信号模型

一个规模为N的种群由S个强者类型的个体和W个弱者类型的个体组成，两种类型的个体发送信号的单位成本分别为c和C（$C>c$）。发送最大信号的所有个体分享利益B。任何大小为$M \geq \frac{B}{SC}$的信号都能够将强者类型分离开来。如果$CW \geq cN$，那么所有强者类型都会分离开来。如若不然，就存在部分混同均衡，其中一部分弱者类型的个体也会发送信号。[2]

这个模型可以解释，为什么昂贵的手表和珠宝能够作为财富的信号。一个人的房子或汽车也标志着其拥有的财富，但是人们无法随时随身携带房屋和汽车。衣服也可以发出财富信号，但是却可能无法创造分离。只要花上几百美元，任何一个人可以穿得像一个拥有大量财富的人。但由于成本很高，手表和珠宝作为财富信号会更加有效。一个穷人或中产阶级人士买不起售价一万美元的手表。戴这样的手表，足以证明自己拥有可观的财富。通过发送这样的信号能够获得的好处可能是，人们会更尊重他，假设人们认为财富在某种程度上与一个人的重要性相关的话（尽管有人可能会质疑这种推断）。

信号的用途和价值

信号能够把隐藏的属性突显出来。我们的行动标志着我们的健康、财富、智慧和慷慨。我们的一些行动产生了作为副产品的信号。一个纯粹是因为对长跑有兴趣而参加马拉松比赛的人，可能会传递出身体健康

和做事专注投入的信号，尽管这可能这不是他的本意。信号模型为解释几乎任何行动提供了另一种视角。但是，一个人选择参加某种活动、学习掌握某种技能、购买某个商品，到底是完全出于个人兴趣，还是在发出某种信号呢？我们也许无法分辨。

例如，信号模型为大学文凭的价值提供了另一种解释。关于收入的大量数据表明，大学毕业生的工资显著高于没有接受过大学教育的人。我们可以推断，较高的工资源于大学期间获得的技能和知识。同时，相关数据还表明，数学和科学专业的大学毕业生工资更高。由此可以推断，在这些专业中学到的技能具有更大的经济价值。然而，如果认真观察一下人们所实际从事的工作，我们可能会发现很少有人在工作中需要使用微积分。而且，求职者在接受面试时，几乎从来没有人被问到过余弦函数的导数怎么求、玻意耳定律怎么解释。有鉴于此，我们可以推断大学学位，特别是科学和数学学位，代表了一个人获取知识能力的信号。毕业生获得的较高薪酬完全取决于学位的信号价值，而不是毕业生在大学期间所学到的知识。[3]

可以考虑一下成为一名医生所必须发送出去的信号。医学专业的学生必须通过物理、有机化学和微积分等课程的考试。但是，医生看病是否使用微积分？医生在为你诊疗耳朵和鼻子的时候，会先在他的记事本上写出一个微积分方程吗？当然不是这样。在很大程度上，微积分知识与医生执业可能完全无关，但是它可能是医生掌握知识体系能力的良好信号。如果真的是这样，那么即便与从事的职业几乎没有任何直接相关性，通过微积分考试也会成为医生的一个有用的信号。

只要有可能，任何人在构建信号的时候，都更愿意在生成信号的同时也能掌握有用的技能。例如，事实证明，要成为一名成功的医生，记忆能力是很重要的。为了传递能够证明记忆能力的信号，面试者可能会要求申请人背出每个国家的首都和货币。成功地通过这项考核，能够证明申请人确实有很强的记忆能力，但是所记忆的这些内容对成为一名好医生并无意义。当你觉得自己的肠胃非常不舒服，匆忙赶到急诊室时，你并不会在乎给你看病的医生是否知道布拉迪斯拉发是斯洛伐克的首都，你只是希望那个医生对消化系统的各个部分都了如指掌。出于这个原因，医疗委员会要求医生通过解剖学考试。通过解剖学考试能够证明一个人的记忆能力，而且记住身体的各个部位也确实是有用的。因此，通过解剖学考试是一个功能性信号（functional signal）。

小结

信号模型的应用范围非常广泛。如前所述，雄孔雀的美丽尾羽是它“身体健康”的信号。众所周知，雄孔雀装饰性极强的扇形尾羽几乎没有任何功能性价值，事实上，这种夸张的尾羽不但无用，而且可能还会给它们带来糟糕的结果。雄孔雀如果选择发展更强壮的爪子，有用性要高得多。但是，强壮的爪子很难让雌孔雀在很远的地方就注意到，这一点比尾羽差得太多了，因此尾羽在演化过程中胜出了。[4]雄性果蝇的彩色尾部也具有与雄孔雀的尾羽类似的功能，蚱蜢和鸟类的鸣叫声也是如此。啁啾需要付出可观的能量，只有吃饱了的蚱蜢才可以花时间啁啾而不用去忙着寻找食物。因此，啁啾声可以起到信号的作用。

在人类社会中，不同的文化会通过不同的行动来表明健康状况。人类学家区分出了三种形式的昂贵信号：无条件的慷慨（unconditional generosity），浪费性的维生方式（wasteful subsistence behavior）、精美的传统手工艺制品。[5]“夸富宴”是居住在太平洋西北地区的印第安土著居民举行的一种仪式，这可能是发送这种慷慨信号的最为突出的一个例子。为了庆祝一个事件，比如一个成员出生或去世时，酋长会送出大量的财富，甚至直接毁坏财富，并对其他酋长提出挑战，要求他们也做出同样的行为。其他酋长如果做不到，就会失去声望。将自己的财物赠予他人，还可以说是有利于社会，但是将它们毁坏无疑是极大的浪费。

事实上，当人们（通常是男性）在预期收益比采集种子或浆果更低时，仍然坚持远行狩猎时，就已经采取了浪费性的维生方式了。男人这样做是因为他们能够希望获得额外的尊重。狩猎成功，猎人就发出了说明他力量和勇气的信号，这在其他环境中也可能很有用。作为一名成功的浆果采集者，能够发送自己拥有良好的视力和耐心的信号，这些当然也是有用的个人特点，这一点毫无疑问，但是在很多方面的预测性能不如狩猎技能好。对生活在澳大利亚北部一个群岛上的梅里亚姆人（Meriam）的一项研究表明，平均来说，作为海龟猎手的男性居民，在50岁时存活的后代人数是其他同龄男性居民的两倍多。[6]

复杂精美的传统工艺品制作需要付出非常多的时间和资源。当然，这种活动也可能生产出有用的物品，如地毯。但是，大多数传统工艺品都是没有太大实用价值的礼仪性物品。一些人类学家将这类传统工艺品的制作解释为信号的发送。创作生产这些物品的意义，不依赖于它们能够实现的功能，因这它们具有重要的文化意义。

很多广告也可以解释为昂贵的信号。例如，购买昂贵的超级碗总决赛的商业广告位，可以说是在发送关于自己产品的“合法性”信号。因为这意味着企业相信消费者会非常喜欢自己的产品，从而可以赚回足以覆盖广告成本的利润。想象一下，假设现在有两家企业分别推出了一款新咖啡机。第一家企业知道自己开发了一个“伟大”的产品。而第二家企业则知道，尽管自己的工程师付出了最大的努力，但是这个产品仍然可能会故障频出导致消费者大量投诉。第二家企业预计将会出现20%的退货率。

每一年，都会有数百万人购买咖啡机。如果不做广告，那么这两家企业可能会平分市场。假设生产出了更好产品的企业决定投入200万美元来宣传自己产品的质量。这家企业预计，在广告攻势下，早期购买者都会购买自己的产品，并且从长远来看，这又会导致更大的销量。这家企业的决策者的脑袋中，可能有一个波利亚瓮模型。相比之下，生产质量较差产品的另一家企业则不会花钱做这样的广告，因为它预计自己的产品应该不会非常畅销。花大钱来表明产品的质量，这种行为有时被称为“烧钱”（burning money）。就像雄孔雀用尾羽吸引到了潜在的配偶一样，企业通过“烧钱”吸引了消费者。

在所有这些情况下，发送信号都要付出成本。那些发出信号的人会发现，信号的成本，会导致他们更大的财富、能力以及慷慨个性被他人识别所能带来的好处而有所减少。此外，发送信号所耗费的时间和精力也可以被认为是一种机会成本：如果把这些资源用在其他用途上，可能会产生更大的社会盈余。例如，一个年轻人可能会花费数小时去决定穿什么衣服，以便表明自己的“社会意识”；或者，一个高中生可能会将大量时间和精力投入到某种“非生产性的”活动中去，因为他相信这样能够提高他被精英大学录取的机会。

为了减少发送信号的社会成本，我们应努力使信号尽可能有效地发挥作用。例如，最好是让年轻人通过参加团体性的运动来证明自己身体健康和勇敢，通过这种运动，他们能够学会体育精神和尊重集体利益，而不要让他们冒着生命危险从飞驰的摩托车上跳下以证明自己的勇敢。最好是要求医生记住人体解剖图谱，而不是考他们记得多少句《魔戒》的精灵语。

尽可能地多尝试吧。浪费性的信号肯定会继续存在。我们的挑战是利用模型，特别是机制设计工具，来构建制度和协议，以保证发送出去的信号确实携带了充分的信息。

26　学习模型

一个人可以养成的最重要的习惯就是对继续学习的渴望。

约翰·杜威（John Dewey）

本章研究个体学习模型和社会学习模型，我们会在两种情况下应用它们。第一种情况，如何学会在一个备选方案集合上做出最优选择。在这种情况下，个体学习和社会学习将会汇聚到最优选择上，而学习规则的不同只能影响收敛速度。第二种情况，如何在博弈中应用学习规则来采取适当的行动。在博弈中，某个行动的收益取决于本人和其他博弈参与者的行动。在这两种情况下，学习规则都更有利于规避风险的均衡结果而非有效率的均衡结果。我们还发现，个体学习并不一定会产生与社会学习相同的结果，而且任何一种学习都不可能在所有环境下都比另一种学习表现得更好。

这些发现为我们的主张——采用多模型方法来表征行为，提供了有力的支持。学习模型介于理性选择模型与基于规则的模型之间。理性选择模型假设人们会审慎考虑所处的环境和要完成的博弈，然后采取最优行动；基于规则的模型则直接根据规则来指定行动。学习模型假设人们会遵循规则，但是，正是这些规则使行为能够发生改变。在某些情况下，行为会趋向最优行为。在这些情况下，学习模型可以用来证明假设人们会采取最优行动的合理性。但是，学习模型也不一定会收敛到均衡，它们也可能生成循环或复杂的动态。而且，如果学习模型确实收敛了，它们可能会有比其他模型更多的均衡可以选择。

本章的内容安排如下。我们首先描述强化学习模型，并将这种模型应用于如何选择最优备选方案的问题。强化学习模型通过更高的奖励来强化行动。随着时间的推移，学习者会学会只采取最优行动。这是一个基准模型，非常适合研究学习模型。它与实验数据也拟合得相当好，而且不仅仅适用于人类。海蛞蝓、鸽子和老鼠，都会强化成功的行动。相比之下，强化学习模型也许更适用于海蛞蝓，它只有不到2万个神经元，而不那么适用于拥有超过850亿个神经元的人类。如此巨大的脑容量使人类能够在学习时考虑反事实，而这种现象是强化学习模型无法考虑的。

然后，我们介绍社会学习模型。在社会学习模型中，个体能够从自

己的选择和他人的选择中学习。个体会复制最流行的或表现高于平均水平的行动或策略。社会学习假设行为者能够观察或沟通。有些物种是通过所谓的共识主动性（stigmergy）来实现社会学习的：成功的行动会留下其他个体可以追随的痕迹或残留物。例如，当山羊在群山间走动时，会留下被踩踏的草，从而强化了通往水或食物的路径。

接着，我们将这两种类型的学习模型应用于博弈分析。如前所述，博弈给出了一个更加复杂的学习环境。同样的行动，可能会在这一个时期内带来高收益，在下一个时期内却产生低回报。正如人们通常可以预料到的那样，我们发现社会学习模型和个体学习模型都不一定会收敛到有效的均衡，而且它们也可能会产生不同的结果。最后，我们讨论了一些更加复杂的学习规则。[1]

个体学习模型：强化学习模型

在强化学习中，个体要根据各个行动的不同权重来选择行动。权重较大的行动比权重较小的行动更经常被选中。分配给某个行动的权重取决于这个个体在过去采取这个行动时所获得的奖励（收益）。这种高回报收益的强化可以导致个体选择更好的行动。在这里，我们要探讨的问题是，强化学习是不是会收敛为只选择具有最高奖励（收益）的那个备选方案。

乍一看，只选择最有价值的那个备选方案似乎是一个非常容易完成的微不足道的任务。如果奖励是完全以数值形式来表示的，例如金钱的数额或时间的长短，那么我们有理由相信人们会选择最好的那个备选方案。在第4章中，我们就是用这种思路来说明一个在洛杉矶工作的人在选择通勤路线时会选择最短的路线。

但是，如果奖励没有采用数值形式（通常情况下都是如此），人们就必须依赖自己的记忆。我们在一家韩国餐厅吃过一次午餐，发现那里的泡菜很美味，所以我们更有可能再次光顾那家餐厅。星期一，我们在跑步前一小时吃了燕麦饼干，结果发现我们连续跑上10千米都不觉得累。如果星期三，我们又在跑步前吃了燕麦饼干并且步履如飞，我们就会加大这个行动（跑步前吃点燕麦饼干）的权重，因为我们已经知道燕麦饼干可以改善跑步成绩。

除了人类之外，其他物种也会这样做。早期研究学习的心理学家爱德华·桑代克（Edward Thorndike）设计了一个经典实验。在这个实验

中，一只通过拉动杠杆逃离了箱子的猫得到了奖励。在回到箱子中之后，这只猫在几秒钟内就再次拉动了杠杆。桑代克得到的数据表明，猫会持续进行尝试。他发现猫（以及人）在奖励增大时学习得更快。他将这个规律称为效果律（law of effect）。[2]桑代克的这个发现是有神经解剖学基础的。重复一个行动会构建出一个神经通路，而这个神经通路在未来会引发相同的行为。桑代克还发现，更出人意料的奖励，也就是远远超出过去奖励水平或预期奖励水平的奖励，会使人们学习得更快，他把这个规律称为惊奇律（law of surprise）。[3]

在强化学习模型中，分配给一个所选备选方案的权重，是根据该备选方案在何种程度上超过了预期，即“渴望水平”（aspiration level）来进行调整的。这样的模型，既考虑了效果律，也就是会采取那些能够更经常地产生更高回报的行动，也考虑了惊奇律，也就是对某个备选方案赋予的权重取决于它所带来的奖励超过了渴望水平的程度。[4]

强化学习模型

假设一个由N个备选方案组成的集合$\{A, B, C, D, \ldots, N\}$、与各备选方案对应的奖励的集合$\{\pi(A), \pi(B), \pi(C), \pi(D), \ldots, \pi(N)\}$，以及一个严格为正的权重的集合$\{w(A), w(B), w(C), w(D), \ldots, w(N)\}$。那么，选择备选方案$K$的概率如下：

$$P(K) = \frac{w(K)}{w(A) + w(B) + w(C) + w(D) + \cdots + w(N)}$$

在选中了备选方案K之后，$w(K)$会增大$\gamma \times P(K) \times (\pi(K) - A)$，其中$\gamma > 0$等于调整速率（rate of adjustment），$A < \max_K \pi(K)$等于渴望水平。[5]

这里需要注意的是，渴望水平必须设定为低于至少一个备选方案的奖励水平。否则，被选中的任何一个备选方案在未来再次被选中的可能性会很低，而且所有备选方案的权重都会收敛到零。不难证明，如果渴望水平低于至少一个备选方案的奖励水平，那么最终几乎所有权重都会被赋予在最优备选方案上。之所以会发生这种情况，是因为每选择一次最优备选方案，权重的增加幅度都会最大，从而给这个备选方案创造了更强的强化。

即便我们将渴望水平设定为低于任何一个备选方案的奖励水平，这种情况也必定会发生。在将渴望水平设定为低于任何一个备选方案的奖励水平时，每种备选方案被选中时权重都会有所增加；因此，这个模型可以用来刻画习惯形成：之所以更频繁地做某件事情，只是因为我们在过去已经做过这件事情。而且，即便将渴望水平设定得很低，会带来最高水平奖励的那个备选方案也会以最快的速度增加权重，因此从长远来看，最优备选方案将会胜出。但是，要收敛到最优备选方案上，所需的时间可能会很长。另外，当我们增加了更多的备选方案时，收敛时间也会变长。

为了避免这些问题，我们可以构造内在愿望（endogenous aspiration）。为此，我们修改上面的模型，将渴望水平设置为平均奖励，从而让它随时间推移而不断调整。想象一下，假设父母试图确定自己的孩子到底是更喜欢苹果薄饼还是香蕉薄饼。设定选择苹果薄饼的奖励为20，选择香蕉薄饼的奖励为10，并将两个备选方案的初始权重都设置为50，将调整速率设定为1，并将渴望水平设定为5。假设父母在第一天准备的是香蕉薄饼，这样香蕉薄饼的权重将增加到55。假设父母在第二天也准备了香蕉薄饼，那么10的奖励等于新的渴望水平，香蕉薄饼的权重不会改变。

假设父母在第三天准备了苹果薄饼。这会带来20的奖励，超出了渴望水平。这会使苹果薄饼的权重增加到60，从而使苹果薄饼变成了更可能被选中的备选方案。更高的奖励也提高了平均收益，因而也使渴望水平上升到了10以上。因此，如果父母再一次准备香蕉薄饼，香蕉薄饼的权重就会减少，因为香蕉薄饼的奖励水平已经低于新的渴望水平了。也就是说，强化学习将收敛为只会选择苹果薄饼。

我们很容易就可以证明，强化学习将趋向于以概率1选择最优备选方案。这个结论意味着，与所有其他备选方案的权重相比，最优备选方案的权重将会变得任意大。

强化学习的效果

在学会选择最优备选方案模型的框架中，当渴望水平被设定为等于平均获得的奖励时，强化学习（最终）几乎总是会选择最优备选方案。

社会学习模型：复制者动态

强化学习假设个体是孤立采取行动的。但是，人们也会通过观察他人来学习。社会学习模型假设个体能够观察到他人的行动和奖励，这可以加快学习速度。现在学界研究得最充分的社会学习模型是复制者动态（replicator dynamics），它假设采取某个行动的概率取决于该行动的奖励和它的受欢迎程度。我们可以将前者称为奖励效应（reward effect），把后者称为从众效应（conformity effect）。[6]在大多数情况下，复制者动态模型都要假定一个无限种群。在这个假设的基础上，我们可以将所采取的行动描述为各种备选方案之间的概率分布。在标准的复制者动态模型中，时间是不连续的，所以我们可以通过概率分布的变化来刻画学习。

复制者动态

假设一个由N个备选方案组成的集合$\{A, B, C, D, \ldots, N\}$、与各备选方案对应的奖励的集合$\{\pi(A), \pi(B), \pi(C), \pi(D), \ldots, \pi(N)\}$。在时间$t$，一个种群的行动可以用这$N$个备选方案上的概率分布来描述：$(P_t(A), P_t(B), \ldots, P_t(N))$。且这个概率分布随如下复制者动态方程而变化：

$$P_{t+1}(K) = P_t(K) \times \left(\frac{\pi(K)}{\bar{\pi}_t}\right)$$

其中，$\bar{\pi}_t$等于第t期中的平均奖励。

考虑这样一个例子，父母可以选择准备苹果薄饼、香蕉薄饼和巧克力薄饼。假设所有的孩子都有同样的偏好，再假设这三种薄饼分别能产生20、10和5的奖励。如果最初有10%的父母制作苹果薄饼、70%的父母制作香蕉薄饼、20%的父母制作巧克力薄饼，那么平均奖励等于10。应用复制者动态方程，在第2期中选择三个备选方案中的每一个的概率如表26-1所示：

表26-1　备选方案概率

备选方案	π	P_1	$\pi/\bar{\pi}_t$	P_2
苹果薄饼	20	0.1	20/10	0.2
香蕉薄饼	10	0.7	10/10	0.7
巧克力薄饼	5	0.2	5/10	0.1

复制者动态方程告诉我们，在接下来一段时间里，制作苹果薄饼的父母的比例将会增大到原来的两倍。这是因为苹果薄饼的奖励等于平均奖励的两倍。而制作巧克力薄饼的父母将会减少，因为巧克力薄饼的奖励只相当平均水平的一半。最后，制作香蕉薄饼的父母的比例则不会改变，因为香蕉薄饼的奖励恰恰等于平均奖励。结合所有这些变化，我们发现平均奖励增大到了11.5。

如前所述，复制者动态同时包括了从众效应（更受欢迎的备选方案更有可能被复制）以及奖励效应。从长期来看，奖励效应占主导地位，因为高奖励的备选方案总是会与奖励水平成比例增长。在复制者动态中，平均奖励发挥的作用，与强化学习中当渴望水平随着平均奖励水平而调整时、渴望水平所发挥的作用类似。两者之间唯一的区别是，在复制者动态中，我们要计算整个种群的平均奖励，而在强化学习中，渴望水平等于个体的平均奖励。只要种群是一个相当大的样本，这种区别就是很重要的。因此，复制者动态产生的路径依赖要远小于强化学习。

在构建复制者动态时，我们需要假设每个备选方案都已经存在于初始种群中。由于最高奖励的备选方案总是具有高于平均奖励水平的奖励，而且它的比例在每个时期都会增加，因而复制者动态（最终）会收敛到整个种群都选择最优备选方案的结果。[7]因此，在这种学会选择最优的环境中，个体学习和社会学习都会收敛到拥有最高奖励水平的那个备选方案上。但是在博弈中则不一定会这样。

复制者动态能够学会最优行动

> 在学会从一个有限的备选方案集中选择最好的备选方案的过程，无限种群复制者动态几乎总是收敛到整个种群都选择最优备选方案。

博弈中的学习

现在，我们将两种学习模型应用于博弈。[8]请先回想一下，在博弈中，博弈参与者的收益不仅取决于他自己的行动，同时也取决于其他博弈参与者的行动。某个特定行动的收益，例如在囚徒困境中的合作，可能会在一个时期内很高而在下一个时期却很低，这取决于另一个博弈参与者的行动。

我们先从油老虎车博弈（Guzzler Game）开始讨论。这是一个双人

博弈，每个博弈参与者都必须选择是驾驶一辆经济型汽车还是一辆高油耗但很坚固的车。选择高油耗的车总能带来2的收益。当一个博弈参与者选择经济型汽车，另一个博弈参与者也选择经济型汽车时，双方都可以得到3的收益，因为两个司机都会有更好的视野，汽车耗费燃油也更少，而且都不必担心被巨大的耗油量压垮。但是，如果另一位博弈参与者选择了油老虎车，那么驾驶经济型汽车的博弈参与者必须非常注意那个人的行为。为了刻画这种影响，我们假设在这种情况下，驾驶经济型汽车的博弈参与者的收益会降低为零。图26-1中给出了这个收益矩阵。

	开油老虎车	开经济型汽车
开油老虎车	2，2	2，0
开经济型汽车	0，2	3，3

图26-1　油老虎车博弈

这个油老虎博弈有两个纯策略均衡：两个博弈参与者同时选择经济型汽车，或者两个博弈参与者同时选择油老虎车。[9]双方都选择经济型汽车的这个均衡会带来更高的收益，这是这个博弈中的有效均衡。

我们先假设，这两个博弈参与者都会进行强化学习。图26-2给出了4个数值实验的结果，其参数为：每个行动集的初始权重都等于5、渴望水平为零、学习速度（γ）为1/3。在这所有4个数值实验中，两个博弈参与者都学会了选择油老虎车，即低效率的纯策略均衡。为什么会这样？为了分析这种情况发生的原因，只需要看一看收益矩阵即可。选择油老虎车的博弈参与者总能得到2的收益；而选择经济型汽车的博弈参与者则有时会得到3的收益，有时则什么也得不到（收益为零）。根据假设，两个行动在初始人口中出现的概率是相等的。因此，选择经济型汽车的平均收益仅为1.5，而选择油老虎车的平均收益则为2。于是就会有更多的博弈参与者选择油老虎车，而这又使选择经济型汽车的收益进一步下降。

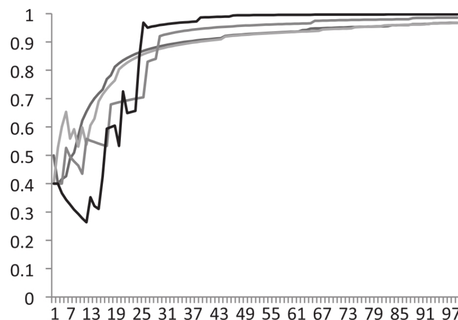


图**26-2**　油老虎博弈中的强化学习：选择油老虎车的概率

接下来，我们将复制者动态应用到这个博弈中来。我们假设初始人口由相同比例的选择油老虎车和经济型汽车的人组成。然后进一步假设每个博弈参与者与每个其他博弈参与者博弈。选择油老虎车的人会获得更高的收益，因为最初选择每种行动的人的数量相等，所以在第二期会有更多的人选择油老虎车。[10]如果再次应用复制者动态方程，那么选择油老虎车的博弈参与者的数量将会进一步上升。持续不断地应用复制者动态方程，最终将导致所有的人都选择油老虎车。图26-3显示的是对一个有100名博弈参与者的油老虎车博弈，运用离散复制者动态进行4个数值实验的结果。

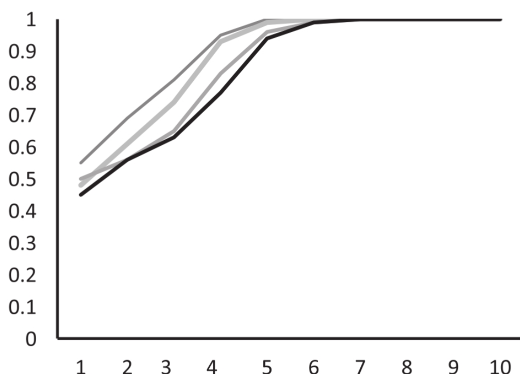


图**26-3**　油老虎博弈中的复制者动态（**100**名博弈参与者）：选择油老虎车的概率

由于假设了有限种群，所以要引入少量随机性，这样采用每个行动的比例可能不完全等于复制者动态方程所给出的比例。在这4个数值实验中，所有博弈参与者全都只在过了7期之后就都选择了油老虎车。收敛之所以如此迅速地发生，原因在于从众效应和奖励效应促使人们在第一期后尽快选择油老虎车。例如，当90%的人选择了油老虎车时，选择

经济型汽车的人的收益将低于选择油老虎车的人的收益的1/6。从众效应极大地放大了奖励效应，使社会学习比个体学习更快。在个体学习中，平均来说要花费100多期才能达到99%的人都选择油老虎车的情况。

在这个博弈中，两个学习规则都收敛到了选择油老虎车上，这是因为当两个行动的可能性相同时，选择油老虎车有更高的收益。这种行动被称为风险主导。两种学习规则都更“青睐”风险主导均衡而不是有效均衡。接下来，我们还要再构造一个博弈模型，在那里，这两个学习规则将收敛于不同的均衡。

慷慨／妒忌博弈

我们要讨论的下一个博弈，慷慨／妒忌博弈（the generous/spiteful game），建立在一个备受关注的人类行为倾向的基础上：我们更加关心的是绝对收益，还是相对收益？假设一个人在以下两个奖金分配方案中选择前者，那么他就是更关心绝对收益：（1）所有同事都能获得15 000美元的奖金，而他自己只能得到10 000美元；（2）所有同事都只能得到5 000美元的奖金，而他自己却能够得到8 000美元。相反，宁愿得到更少奖金（在上面两个奖金分配方案中选择后者）的人则更关心相对收益。这种更关心相对收益的偏好，在“居心险恶的人与魔法灯”的寓言故事中得到了很好的体现。

居心险恶的人与魔法灯

在一次考古探险中，一个居心险恶的人发现了一盏青铜灯。他擦了一下灯，结果召唤出了一个精灵。精灵说：“我会赐予你一个愿望，因为我是一个仁慈的精灵。我可以给你想要的任何东西！不但如此，对于你认识的每一个人，我都将给予他们给你的两倍。”这个居心险恶的男人仔细想了一会，然后抓起了一根棍子，递给那个精灵，说：“好吧。现在请你抠出我的一只眼睛吧。”

这个居心险恶的人所采取的这个行动，在带给他一个很低的绝对收益的同时，又给了他一个高的相对收益。[11]在外交事务中也会出现类似的紧张局势。新自由主义者认为，每个国家都希望最大化自己的绝对利益，这可以用军事力量、经济繁荣和国内稳定来衡量。但是另外一些人，他们通常被称为新现实主义者，却认为各国更重视的是相对利益，

即一个国家宁可只能获得较低的绝对利益，但是一定要比自己的敌人更强。在冷战高潮期间，新现实主义者肯尼思·华尔兹（Kenneth Waltz）这样写道："各国首先关心的不是最大化自己的力量，而是要维持他们在国际体系中的地位。"[12]

我们可以将绝对收益与相对收益之间的这种潜在冲突嵌入到一个N人博弈中。在这个博弈中，存在两类行为，一种行为是"慷慨大度"的，它在增加行为者自己绝对收益的同时，也会增加其他人的收益；另一种行为则是"妒忌狭隘"的，它只会增加行为者自己的收益。这个博弈不同于集体行动博弈。在集体行动博弈中，慷慨大度是要付出成本的。[13]在这个模型中，慷慨的行动是占优策略：无论其他博弈参与者采取的行动是什么，选择慷慨行动的博弈参与者都能获得更高的收益。然而，选择妒忌行动的博弈参与者得到的收益却比选择慷慨行动的博弈参与者更高。

这些陈述，乍一看似乎是自相矛盾的，其实不然。通过选择慷慨的行动，这个博弈参与者能够将自己的绝对收益提高3，而且同时也会将所有其他博弈参与者的收益提高2。而选择妒忌行动的博弈参与者则只能使自己的收益提高2，而且不能提高其他博弈参与者的收益。因此，每个博弈参与者都能通过选择慷慨的行动来提高自己的收益。相反，当一个博弈参与者选择妒忌的行动时，他反而会降低自己本来可以得到的收益，而且（这是关键假设）会使其他人的收益下降得更多。

慷慨／妒忌博弈

在这个博弈中，有N个博弈参与者，每一个博弈参与者都要选择：是慷慨（G），还是妒忌（S）。

收益（G，N_G）$=1+2\times N_G$

收益（S，N_G）$=2+2\times N_G$

如果我们在慷慨／妒忌博弈中应用强化学习规则，那么博弈参与者们都会学会做一个慷慨的人。为什么会这样？要理解这一点，不妨假设博弈参与者处于几乎已经收敛到均衡的状态了，N_G个博弈参与者都选择了慷慨的行动。这时，如果有一个妒忌的博弈参与者，那么他可以获得$2+2\times N_G$的收益。这将成为他的渴望水平。如果他选择的是G（哪怕发生概率很小），那么他的收益为$1+2\times(N_G+1)=3+2\times N_G$，这比他的渴望水平还要高。因此，他将更有可能变成慷慨的。不断应用这种逻

辑，我们就会发现所有的博弈参与者都将学会慷慨。

但是，如果我们在慷慨/妒忌博弈中应用复制者动态，人们就会学会做一个妒忌的人。只要仔细观察一下复制者动态方程，就可以得出这个结论。在每一期，选择妒忌行动的那些博弈参与者所获得的收益都要高于选择慷慨行动的博弈参与者。因此，在每一期，选择妒忌行动的博弈参与者的比例都会上升。

这些结果突显了个体学习与社会学习之间的关键差异。个体学习会引导人们选择更好的行动，因此人们会学会采取占优行动（如果占优行动存在的话）。而社会学习则引导人们选择相对于其他行为来说表现更好的行动。在大多数情况下，这些行动通常也会产生更高的收益。但是，在慷慨/妒忌博弈中却并不是这样。在这种模型中，妒忌行动的平均收益更高，但是占优的却是慷慨行动。特别需要注意的是，我们的分析得出了一个悖论性质的结果：如果人们进行的是个体学习，那么他们就能够学会做一个慷慨的人——比通过社会学习能够学会的更加慷慨。之所以会出现这种情况，是因为在社会学习中，博弈参与者会复制表现相对较好的博弈参与者的行动。

现在考虑一下之前的观点：我们可以将复制者动态视为一种适应性规则，或者视为发生在若干固定规则之间的选择。如果假设了后者，那么我们的模型就意味着选择可能有利于妒忌这种类型。（自然）选择不一定会导致合作。这个结果与我们在研究重复囚徒困境博弈时发现的结果背道而驰——在那里，重复导致了合作。我们在那里考虑的是重复博弈，并允许更复杂的策略。

将不同的学习模型组合起来

我们已经看到了个体学习模型和社会学习模型都能在一组固定的备选方案中找到最优解决方案，但当把它们应用于博弈时，也可能产生不同的结果。缺乏协议也可以是一种力量。不妨想象一个由所有可能的博弈组成的巨大集合。再想象一个由所有学习模型组成的巨大集合。我们可以将第二个集合中的每个学习模型应用于第一个集合中的每个博弈，并评估它们的表现。然后我们可以将这个包括了所有博弈的集合划分为两个集合：学习规则产生了有效结果的博弈的集合，以及不能得到有效结果的博弈的集合。我们还可以考察实验数据并评估每个学习规则在作为对实际行为的预测器时表现如何。毫无疑问，这种研究将会揭示一些

我们未曾预料过的东西。每个学习规则都会在某些博弈中带来有效的结果，但是对其他一些博弈则不能。同时，每个学习规则本身在准确描述了行动的不同情况下也会有所不同。因此，我们提倡多模型思维。

在本章中，我们介绍了两个标准的学习模型。每一个模型都只包括了少数几个会变化的组成部分。我们的目标是对这些令人兴奋的文献给出一个适中的介绍。只要往这两个标准模型中加入更多细节，就能够更好地拟合实验数据和经验数据。请读者回想一下，在强化学习模型中，个体会根据一个已采用过的行动或备选方案的奖励（或收益）是否超过了渴望水平来加大或减少该行动或备选方案的权重。个体不会增加从未采取过的行动的权重：在强化学习模型中，我们不会因为假想采取了某个行动会带来很高的收益，而提高采取这种行动的概率。

在所有情况下，这种假设都没有意义。假设一个员工决定在休假时不带手机。当他去度假时，他的老板打电话要他解决一个重要问题，这个员工没有接到这个电话，并因此错过了一个升职机会。在强化学习模型中，员工不会给“度假时带手机”这个行动赋予更大的权重。有鉴于此，人们对标准模型进行了修正，提出了罗斯-伊雷夫学习模型（Roth-Erev learning model），让未被选择过的备选方案也可以根据其假想的收益来获得权重。在这个例子中，这个员工将会给“度假时带手机”赋予更大的权重。

这个修正导致了基于信念的学习规则。未被选择过的备选方案权重的增加量可以通过一个实验参数来确定。实验参数越高，人们对其他人行为的影响的考虑就越多，对那些行为赋予的权重也增加得越多。经济学家埃尔文·罗斯（Alvin Roth）和伊多·伊雷夫（Ido Erev）还考虑到，其他博弈参与者也在学习，他们的策略也可能在发生变化，因此还对过去进行了贴现处理。[14]

这些额外的假设具有其直观意义，并且都得到了经验证据的支持，但是它们并不适合于所有情况。如果回到前面举过的制作薄饼的例子，那么第一个假设意味着在父母制作好了香蕉薄饼之后，还要赋予制作苹果薄饼的备选方案额外的权重，而且该权重要与苹果薄饼的收益成比例。只有当父母知道苹果薄饼的收益时，这样的假设才是有意义的。但是，只有当人们能够观察到或凭直觉感知未被选择行动的收益时，才会出现这样的情况。

第二个修正模型来自行为经济学家科林·凯莫勒（Colin Camerer）和何（Ho）。他们构建了一个通用的函数形式，把强化学习和基于信念

的学习都作为特殊情况包括了进去。这个函数的关键是一个可以用数据拟合的、确定每种类型学习规则相对强度的参数。[15]将多个模型组合在一起，正是我们学习掌握许多模型的一个重要动机。也就是说，由于参数的增加，组合模型必定能够导致更好的拟合。即便考虑到了参数增多这个因数，凯莫勒和何的模型也能给出更好的预测和更深刻的解释。

对学习建模带来了一些挑战。在一个模型设置中运行良好的学习规则可能完全无法适用其他情况。此外，人们学习的东西可能取决于他们最初的信念，因此两个人可能在同一个环境中以不同方式学习，同一个人也可能在不同的环境中以不同的方式学习。即便我们真的构建出了一个准确的学习模型，也会遇到可利用性原则（exploitability principle）带来的难题：如果一个模型解释了人们如何学习，那么其他模型就可以应用这个模型来预测相关知识，并在某些情况下利用该知识。这样一来，人们就可能会学会如何不会被利用，从而使我们原来的学习模型不再准确。在本书前面的章节中，当我们讨论卢卡斯批判和对有效市场假设的分析时，我们就已经遇到过这种现象了。我们不一定能得出结论说那是因为人们会了解到他们在最优化，然而，学习毕竟倾向于淘汰不良行为、从而有利于更好的行为。

文化能否压倒战略

我们现在将传染模型和学习模型结合起来，以便剖析组织理论中由来已久的一个理论观点：文化压倒战略。[16]简而言之，这个观点声称，改变行为的战略激励终将归于失败。理论组织家强调，文化——即现有的既定规则和信念的力量实在太强大了。经济学家的观点则相反：推动行为的，只能是激励。

为了将这些相对立的谚语式诊断转变成条件逻辑判断，我们首先必须应用网络传染模型的一个变体。在这个模型中，经理，或者也可能是CEO，宣布了一个新战略，并给出了推动变革所能带来的好处的多项证据。这位经理或CEO甚至可能会对组织的核心原则加以重新界定，以便反映这种新行为的要求。然后，组织中的其他个体决定是否采取这种行为，这取决于经理或CEO对其战略的说服力有多大。一开始，只有一部分人执行这个计划。当他们在工作网络中与他人互动时，就会热情洋溢地传播新战略。当然新战略也会面临挑战，会有一种反向的力量拉动人们不去采用新战略。有三个特征决定了新策略能否顺利展开：接触率（$P_{contact}$）、扩散率

（P_{spread}）和放弃率（$P_{recover}$），它们很自然地映射到了基本再生数中的参数R_0上，即：

$$R_0=\frac{P_{spread}\times P_{contact}}{P_{recover}}$$

如果再加入存在超级传播者的可能性，就可以得出这样的结论：只要如下三个条件中任何一个条件成立，文化就会压倒战略，否则，战略就能压倒文化。这三个条件分别是：如果人们不相信新战略，如果人们很快就放弃了新战略，如果新战略的拥护者相互之间的连通性不够好。

我们的第二个模型是，将复制者动态方程应用于这个用来表征员工之间互动的文化战略博弈。我们可以将员工的不同选择用博弈论的语言分别表示为文化行动（做他们目前所做的事情）和创新的战略行动。我们还假设，经理或CEO已经确定了收益结构，如果两个博弈参与者都选择创新的话，他们都能获得更高的收益；但是，如果只有一个博弈参与者选择创新，那么他的收益将会减少。

	文化	创新
文化	200，200	220，180
创新	180，220	300，300

文化 / 战略博弈

这个博弈有两个严格的纯策略纳什均衡：一个是两个博弈参与者都创新（战略胜过了文化），另一个是两个博弈参与者都不创新（文化胜过了战略）。乍一看，经理或CEO似乎已经给出了足够大的激励，能够保证员工会选择创新的行动。但是通过分析，我们发现，经理或CEO必须动员起足够多的初始支持者才能使创新成为现实。如果一开始就支持新战略的人的比例没有超过20%，那么文化就会胜过战略。如果要增加创新战略的收益，那初始支持者的比例可能会更低，但仍然会产生有效的结果。[17]

这两个模型表明，字面上相反的两个谚语“文化压倒战略”和“人们会对激励做出反应”都是正确的。根据第一种模型，具有很高人格

魅力的CEO可以制订能够胜过文化的新战略。根据第二种模型，文化能胜过“弱激励”，但是不能胜过“强激励”。

27 多臂老虎机问题

有一件事我确实特别擅长，那就是将网球击过网，打在界内。在这件事情上，我是最棒的。

塞雷娜·威廉姆斯（Serena Williams）

在本章中，我们在如何学会选择最优备选方案的学习模型中加入不确定性，从而生成了一类被称为“多臂老虎机问题”（multi-armed bandit problems）的模型。在一个多臂老虎机问题中，不同备选方案的奖励源于一个分布，而不是固定的金额。多臂老虎机问题模型适用于各种各样的现实环境。在收益不确定的行动之间进行的任何选择，无论是药物试验，还是对树立广告牌位置的选择、技术路线的选择，抑或是要不要允许在教室中使用笔记本电脑的决定，都可以建模为多臂老虎机问题。当然，如何选择一个可以出人头地的职业，也可以用多臂老虎机问题模型来建模。[1]

在面对一个多臂老虎机问题时，人们必须对各个备选方案多加尝试，以便通过这种学习过程来了解收益的分布。多臂老虎机问题的这个特征，也导致我们必须在探索（寻找最佳备选方案）和利用（选择迄今为止表现最佳的备选方案）之间善加权衡。在探索与利用权衡中找到最优平衡点，需要非常精妙复杂的规则和行为。[2]

本章的主体内容分为两个部分，最后对模型的应用价值进行了讨论。在本章的第一部分，我们描述了一类特殊的伯努利多臂老虎机问题，其中每个备选方案都是一个伯诺利瓮（瓮中灰球和白色球的比例是未知的）。我们描述并比较多种启发式求解方法，然后说明这些解是如何有助于改进药物疗效的比较检验、广告计划和教学策略的。在第二部分中，我们描述了一个更一般的模型，其中收益分布可以采取任何形式，并且决策者对其类型有一个先验分布。我们还阐明了如何求解确定最优选择的吉廷斯指数（Gittins index）。

伯努利多臂老虎机问题

我们从一类特殊的多臂老虎机问题开始讨论。在这类多臂老虎机问题中，每个备选方案都能以固定的概率产生成功的结果。因此，这类多

臂老虎机问题相当于在一系列伯努利瓮之间进行选择，且每个瓮都包含着不同比例的灰球和白球。因此，我们将这类多臂老虎机问题称为伯努利多臂老虎机问题，也经常被称为频率问题，因为决策者对分布一无所知。不过，当决策者对各个备选方案进行了多次实验（探索）之后，他会对这些分布有所了解。

伯努利多臂老虎机问题

一个备选方案集{A，B，C，D，…，N}中的每一个备选方案都能够产生一个成功的结果，但是各自的概率{P_A，P_B，P_C，P_D，…，P_N}都是未知的。在每一个时期，决策者选择一个备选方案K，并以概率P_K得到一个成功的结果。

假设一家烟囱清洁公司获得了一批最近购买了房子的人的电话号码，然后打算向他们推销烟囱清洁服务。这家公司测试了三种推销策略。第一种策略是“笃定预约式”（“你好，我打电话来是为安排你家每年一度的烟囱清洁。”）；第二种策略是“关心提问式”（“你好，你知道烟囱堵塞是火灾的最大风险因素吗？”）；第三种策略是“人性感动式”（“你好，我的名字是希尔迪，我已经和我父亲一起打理这家烟囱清扫公司整整14个年头了。”）。

每一种推销策略都有可能成功，但是成功概率在事前是未知的。假设该公司首先尝试的推销策略是“笃定预约式”，但是失败了。然后又尝试了“关心提问式”，结果成功地获得了一个客户；而且，这种策略紧接着又成功了一次。但是，在接下来的3次尝试中，这种策略都失败了。于是，该公司尝试了第三种策略。第三种策略的第一次尝试是成功的，但是接下来却连续失败了4次。这样，在总共进行了11次尝试之后，第二种推销策略的成功率是最高的，但是第一种策略只尝试了一次。于是，决策者面临着在利用（选择最有效的备选方案）或探索（回过头去继续尝试其他两种推销策略以获得更多信息）之间的权衡。医院在不同的外科手术方案之间的选择，制药公司对药物检验的不同方案的权衡，也都会碰到同样的问题。每一种“协议”都有未知的成功概率。

为了进一步深入理解这种探索-利用权衡，我们比较了两种启发式。第一种启发式是取样并择优启发式（sample-then-greedy），即先对每个备选方案都尝试固定的次数M，然后选择具有最高平均收益的备选方案。而在确定尝试次数M大小的时候，我们可以参考伯努利瓮模型和平

方根规则。平均比例的标准差有一个上界$1/2\sqrt{M}$。如果每种备选方案都进行了100次的测试，那么平均比例的标准差将等于5%。如果应用两个标准差规则来识别显著差异，当两个比例相差大约10%时，我们就能够自信地将它们区分开来了。例如，如果一个备选方案在70%的时间内都取得了成功的结果，而另一种方法则只在50%的时间内取得了成功，那么就有95%以上的置信水平相信我们能够选中正确的备选方案。

第二种启发式称为自适应探测率启发式（adaptive exploration rate heuristic）。它的程序是，第一阶段，先让每种备选方案各完成10次试验。第二阶段，再进行总共20次试验，但是试验次数根据各备选方案在第一阶段的成功率按比例分配。例如，如果第一阶段的10次试验中，有一个备选方案成功了6次而另一个只成功了2次，那么第一个备选方案将获得接下来的20次试验中的3/4。到第三个阶段，可以根据成功率的平方确定本阶段要进行的20次试验的分配比例。如果那两个备选方案的成功率仍然与前面一样，那么更好的备选方案在第三阶段的20次试验中将分配到$\frac{(0.6)^2}{[(0.6)^2+(0.2)^2]}$或90%。

以此类推。对于每一个阶段的20次试验，计算分配比例时所用的成功概率的指数可以以某种速率递增。随着时间的推移，通过提高利用率，第二种算法相比第一种算法有所改进。如果一个备选方案的成功概率比另一个备选方案高得多，比如80%对10%，那么第二种算法就不会在第二个备选方案上“浪费”100次试验。另一方面，如果两个备选方案的成功概率非常接近，就要继续进行试验。[3]

对于第一种启发式，取样并择优启发式，如果在使用时过分执着，那么不仅效率低下，而且可能是不道德的。当美国著名外科医生罗伯特·巴特莱特（Robert Bartlett）在测试人工肺时，发现它的成功率远远超过了其他备选方案。既然人工肺的表现已经如此优异了，那么继续测试其他备选方案就会导致不必要的死亡。于是巴特莱特停止测试其他备选方案，让每个患者都使用上了人工肺。事实上，可以证明这是一个最优规则：如果某个备选方案总能取得成功，那么就继续选择这个备选方案。增加实验可能没有任何价值，因为没有其他备选方案能够表现得更好。

贝叶斯多臂老虎机问题

在贝叶斯多臂老虎机问题中，决策者对各备选方案的收益分布有先验信念。考虑到这些先验信念，我们可以对探索与利用之间的上述权衡进行定量分析，并（至少在理论上可以）在每个时期都做出最优决策。然而，即便是对于最简单的多臂老虎机问题，要确定最优行动也需要进行大量的计算。在真实世界的实际应用中，要精确计算出结果是不可行的。因此决策者通常都会利用近似方法。

贝叶斯多臂老虎机问题

给定备选方案集$\{A, B, C, D, \ldots, N\}$，以及对应的收益分布$\{f(A), f(B), f(C), f(D), \ldots, f(N)\}$。决策者对每个分布都有先验信念。在每一期，决策者选择一个备选方案，并获得收益，并根据收益计算出新的信念。

要确定最优行动，需要经过如下四个步骤。首先，要计算出每个备选方案的即时期望收益。其次，对于每个备选方案，都要更新关于收益分布的信念。再次，在得到的关于收益分布的新信念的基础上，根据我们所掌握的信息确定所有后续时期的最优行动。最后，我们将下一期行动的期望收益与未来的最优行动的期望收益相加。最后得到的这个结果就是通常所称的吉廷斯指数。在每一个时期，最优行动的吉廷斯指数都是最大的。

这里需要注意的是，计算指数的过程同时也量化了探索的价值。而且，如果我们尝试某个备选方案，那么吉廷斯指数也不会等于期望收益。相反，吉廷斯指数等于假设根据所掌握的知识采取最优行动时，所有未来收益的总和。但是，计算吉廷斯指数非常困难。下面举一个相对简单的例子。假设备选方案有两个，一个是肯定能带来500美元的安全的备选方案，另一个是有10%概率可以带来1 000美元的有风险的备选方案，在其余90%的时间里，有风险的这个备选方案不会带来任何收益。

为了计算出这个有风险的备选方案的吉廷斯指数，我们首先要问清楚会发生什么：它要么总是收益1 000美元，要么总是没有任何收益。然后再思考每个结果将会怎样影响我们的信念。如果我们知道这个有风险的备选方案会让我们收益1 000美元，那么我们总是会选择它。如果我们知道这个有风险的备选方案没有任何收益，那么我们在未来将总是选择安全的那个备选方案。

由此可见，有风险的这个“臂”的吉廷斯指数对应于每个时期获得1 000美元奖励的概率为10%，以及除了第一期之外的每个时期获得500美元的概率为90%。平均来说，要对备选方案进行多次选择的情况下，这相当于每一期大约550美元。因此，这个有风险的备选方案才是更好的选择。[4]

吉廷斯指数

为了说明如何计算吉廷斯指数，考虑下面这个只有两个备选方案的例子。备选方案A产生的收益抽取自{0，80}，且0和80出现的概率相等。备选方案B在{0，60，120}当中产生一定的收益，而且这三个收益的概率也是相等的。我们假设，决策者试图最大化10个时期内的总奖励。

备选方案**A**：收益等于零的概率为1/2，在出现了这种结果之后，在剩下的全部9期内都会选择备选方案B（备选方案B的期望收益60），这样就得到了540的期望收益（9乘以60）。收益等于80的概率也为1/2，即便出现了这个结果，在第二期的最优选择仍然是选择备选方案B。于是有1/3的概率，备选方案B产生了120的收益，因此总收益等于1 160（80加上9乘以120）。同样，有1/3的概率，备选方案B产生了60的收益，在这种情况下，备选方案A是剩下的所有8期的最优选择，这样产生的总收益等于780（60加上9乘以80）。最后，还有1/3的概率，备选方案B产生了零的收益，在这种情况下，备选方案A是剩下的所有8期的最优选择，这样产生的总收益等于720（9乘以80）。

把上面这些可能性全都考虑进去，可以得出，在第一期，备选方案A的吉廷斯指数如下：

$$\text{Gittins index}_1\,(A)=\frac{1}{2}\times 540+\frac{1}{2}\left(\frac{1}{3}\times 1160+\frac{1}{3}\times 780+\frac{1}{3}\times 720\right)=\frac{2140}{3}$$

备选方案B：有1/3的概率，收益等于120。如果发生了这种情况，那么所有未来时期的最优选择也仍然是备选方案B。因而10个时期的总收益将等于1 200。如果收益等于零（概率为1/3），那么所有未来时期的最优选择都将是备选方案A（备选方案A的期望收益为40），因而，期望总收益将等于360（9乘以40）。如果收益等于60，那么决策者在所有未来时期都应该选择替代方案B，总回报

为600；但是，如果在第二个时期选择了备选方案A，那么有一半时间备选方案A总是产生80的收益，此时总回报为780（60加上9乘以80）；另一半时间它产生零收益，并且所有后续时期的最优选择都将是备选方案B（会产生60的收益），于是得到的总收益为540（9乘60）。由此可知，在第二期中选择备选方案A才是最优选择，这种选择产生的期望收益等于660×（1/2×780＋1/2×540）。

把所有可能性都考虑进去，不难推出，备选方案B在第一期中的吉廷斯指数如下：

$$\text{Gittins index}_1(\text{B})=\frac{1}{3}\times 1200+\frac{1}{3}\times 660+\frac{1}{3}\times 360=\frac{2220}{3}$$

根据这些计算结果，备选方案B是第一期的最优选择。最优长期选择则取决于第一期内学习的结果。如果备选方案B产生了120的结果，那么我们将永远坚持选择备选方案B。

上述的分析表明，在采取行动时，我们更关心的是备选方案能够成为最优选择的概率而不是它的期望收益。此外，如果某个备选方案会产生非常高的收益，我们应该更有可能在将来选择它。相反，如果它只能产生平均收益，那么即便收益水平高于另一种备选方案的期望收益，我们一直坚持这个备选方案的可能性也不会太高。在尝试的早期阶段尤其如此，因为我们希望找到更高收益的备选方案。这些结果在我们讨论的诸多应用中都是成立的。如果采取行动没有风险，也不需要付出高额成本，那么这个模型告诉我们，即使高收益行动的概率很低，我们也要努力去探索它们。

小结

本书一再强调的一个核心要点是，通过学习模型，我们可以做出更好的决策。在对人们在多臂老虎机问题中“应该”做些什么与人们“实际”做了些什么进行一番比较之后，我们可以对这一点有更深刻的理解。绝大多数人在遇到多臂老虎机问题时，都不会去试图计算一个吉廷斯指数。之所以没能这样做，部分原因是他们没有将必要的数据保存下来。例如，直到最近，医生才开始记录各种手术方法的效果，比如不同类型的人工关节的功效，或者不同类型的心脏支架的优缺点，等等。没有这些数据，医生就无法确定哪一种手术方法会带来最高的期望收益。

其实是所有人，都需要足够的数据，这样才能把模型教给我们的东西应用起来。因此，如果你真的想了解晚餐前散步还是晚餐后散步对你的睡眠更有利，你就需要跟踪记录你的睡眠状况，并运用一些相当复杂的启发式去了解哪种散步模式效果更好。初看上去，这可能会显得小题大做，而且要付出的时间精力相当可观，的确如此，但是现在已经好很多了。新技术的不断涌现，使我们能够非常方便地收集有关睡眠模式、脉搏率、体重，甚至情绪好坏的大量数据。

我们每个人都要做出很多与自己的身体健康息息相关的决策，比如什么时候去锻炼，但是绝大多数人都没有去收集必要的数据并计算出吉廷斯指数。但这其实非常重要，关键在于我们是否能够做到这一点，而且如果做到了，我们的睡眠模式、身体健康状态就会得到改善。心理学家塞思·罗伯茨（Seth Roberts）探索了整整12年，结果发现自己每天至少站立8小时才可以改善自己的睡眠状态（尽管他还是睡得更少）。他还发现，迎着早晨的阳光站立，可以减轻他上呼吸道的感染症状。[5]当然，我们一般人可能很难具备他这种用自己身体来做实验的奉献精神。但是，由于不保存数据，也不对相关结果进行比较，我们可能会更容易放任自己不吃早餐或暴饮暴食——尽管我们有更好的选择，比如吃西柚。

在利害关系很大的商业决策、政策制定和医疗决策中，数据更容易收集，应用多臂老虎机问题模型也早就成了一种常见的做法。企业、决策者和非营利组织，都会先对各种备选方案进行探索，然后利用那些表现最好的备选方案。而且在实践中，备选方案往往不会保持固定不变。例如，鼓励参加农业补贴计划的政府邮件可能每一年都会改变，比如将上一年的强健男子的照片换为性感美女的照片，等等。[6]这种类型的连续实验可以通过将在下一章中讨论的模型来刻画，那就是：崎岖景观模型。

用不同模型分析美国总统选举

我们可以应用至少三种模型来分析美国总统选举：空间竞争模型、分类模型和多臂老虎机问题的模型。

空间竞争模型：民主、共和两党的总统候选人在意识形态空间中互相竞争以吸引选民。我们有理由预计，各候选人倾向于温和的中间立场、选情会比较胶着，不同政党候选人获胜的顺序是随机的。除了少数例外情况外，总统选举不会以某一方压倒多数获胜结

束。为了检验美国总统大选获胜者的顺序是不是随机的，我们构建了从1868年到2016年38次总统大选的获胜政党的时间序列。该序列如下（字母R、D分别表示共和党、民主党）：

RRRRDRDRRRRDDRRRDDDDDRRDDRRDRRRDDRRDDR

然后我们可以计算出不同长度的子序列（块）的熵。长度为1的子序列的熵为0.98。长度为4的子序列的熵为3.61。统计检验表明，我们不能否认这个序列是随机的。作为比较，在长度为38的随机序列中，长度为1的子序列的熵为1.0，长度为4的子序列的熵为3.58。

分类模型：如果我们将每个州视为一个类别，同时假设不同州之间存在着异质性，那么空间竞争模型意味着一旦候选人选定了初始位置，某些州就不再具有竞争力了。这个模型的预测是，在少数几个立场温和的州，选举竞争将特别激烈。2012年，奥巴马和罗姆尼都在10个州花掉了自己电视广告预算的96%以上。他们每个人都将广告预算的一半多用于3个温和的州：佛罗里达州、弗吉尼亚州和俄亥俄州。2016年，希拉里·克林顿和特朗普也将一半以上的电视广告预算花在了3个温和的州：佛罗里达州、俄亥俄州和北卡罗来纳州。[7]

多臂老虎机问题的模型（回溯性投票）：选民将更有可能将选票再一次投给有良好执政业绩的那个政党。给绩效好的政党投票，相当于拉一个会带来高收益的杠杆。经济繁荣通常会使竞选连任者受益。有证据表明，当经济表现良好时，选民更有可能投票给执政党的候选人。而且，在执政党内部，现任候选人的影响也大于非现任候选人。[8]

28　崎岖景观模型

当你费心去寻找时，就会发现令人惊奇的事情。

传为萨卡加维亚（Sacagawea）所说

在本章中，我们研究崎岖景观模型。与空间竞争模型和享受竞争模型一样，崎岖景观模型也将一个实体定义为属性的集合。每个属性的集合都映射到一个价值上。崎岖景观模型的目标是修改属性，以构造出一个具有最高价值的实体。这类模型起源于生态学中对演化的研究。现在，崎岖景观模型已经广泛用于探索各种问题的求解方法、研究企业之间的竞争和创新，以及其他领域，这也是我们在本书中要研究的重点。在本章中，我们将应用崎岖景观模型揭示，属性影响的相互依赖性如何使创新变得困难、导致所找到的解决方案呈现出路径依赖性、并且还造成了解决方案本身的多样化。同时，我们也会阐明，许多更困难的问题是怎样通过更加多样化的问题求解方法而得以解决的。

本章由三个部分组成，然后，我们讨论了如何扩展模型以刻画竞争。在第一部分中，我们先描述了一个适合度景观模型（fitness landscape model），然后阐明了怎样将它重新解释为一个关于问题求解和创新的模型。在第二部分中，我们讨论了一维模型中崎岖度的含义。在第三部分中，我们提出了崎岖景观的NK模型，它将一维模型扩展到了任意数量的二元维数。

适合度景观模型

适合度景观模型假设物种拥有能够促进其适合度的特征或性状，我们可以不那么严格地将之定义为繁殖潜力，同时种群中不同成员所拥有的特定性状的数量或程度可能不同。如果用横轴表示性状，用纵轴表示物种的适合度，就可以绘出一张适合度景观的图，其中高海拔点对应高适合度。

举例来说，为了绘制出一张对应于土狼尾巴长度性状的适合度景观的图，我们应该令土狼的所有其他性状都保持不变，而只改变尾巴的长度，并测量尾巴长度的变化对适合度的影响。这就是说，要绘制出这张图，我们必须先了解土狼尾巴为什么有助于提高它的适合度。

假设一只土狼的尾巴有助于土狼在跳跃时保持平衡，而且土狼可以将它作为表示幸福、恐惧或即将发动攻击的信号。我们从横轴的最左侧开始，在那里，尾巴长度为零，这种情况下它不能执行任何一种功能，因此它的适合度为零。随着尾巴长度的增加，维持平衡和传递信号的功能也随之提高。因此，适合度先是随尾巴长度的增加而上升的。

但是，到了某一点上，比如当尾巴长到18英寸时，可能就是有助于土狼保持平衡的最理想长度。如果尾巴变得更长，土狼运动的敏捷度将会下降。不过，更长的尾巴可能还会继续提高它传递信号的价值，因此，长度为20英寸的尾巴可能会产生最大的整体适合度。一旦尾巴的长度超过了20英寸，适合度就会开始下降。结果如图28-1所示，它具有一个单峰。

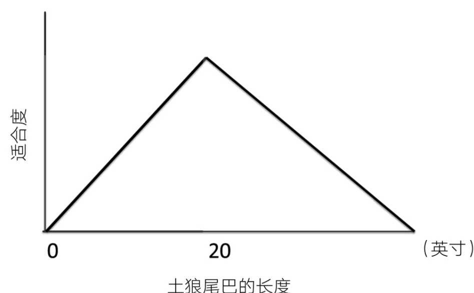


图28-1　富士山景观

这种单峰景观被称为富士山景观。在现实世界中，这样的景观是经常出现的。有富士山景观的问题通常是比较容易解决的问题。我们可以指望演化或学习在遇到这类问题时找到这个山峰。不妨想象一下，假设有一个由尾巴长度各不相同的土狼组成的种群，那么自然选择的压力最终会导致土狼的尾巴长到大约20英寸。拥有这个长度的尾巴的土狼，能够同时将维持身体平衡和传递信号这两个任务完成得最好。因此，这样的土狼的适合度最高，能够留下最多的后代，从而导致更多的土狼拥有20英寸长的尾巴。如果我们认为这是一个优化问题，那么任何一个爬山算法都能找到这个山峰。

我们也可以应用一对多的思维方法，将这个问题重新解释为一个产品设计问题，比如设计一把煤铲。假设我们已经决定了煤铲的长度和形状，只剩下煤铲的大小有待决定。于是，煤铲的面积就是要反映在横轴上的特征。而在纵轴上，我们给出的是一个工人在给定煤铲大小的前提下，一个小时内能够铲煤的数量。

与之前的例子一样，我们仍然从最左侧开始，那一点对应于面积为零的煤铲。面积为零的煤铲其实只是一根棍子。当然，用一根棍子去铲煤是完全无效率的，其价值为零。我们逐渐增大煤铲的面积，先是一茶匙大，接着是一汤匙大，然后是像玩具铲那么大……煤铲就变得越来越有效。在这个过程中，煤铲的适合度图形是向上倾斜的。然后，到了某个点上，当煤铲的面积变得太大之后，要用它去铲煤就变成了一件苦差事。一个小时内可以铲煤的数量就会随煤铲的面积进一步增大而减少。最后，当煤铲面积变得足够大时，将没有任何人能够抬起这把煤铲，因而它的适合度又一次变为零。我们又看到了富士山景观。我们有理由期待肯定能够找到山峰，也就是要设计的煤铲的理想面积。

事实上，将铲子的效率作为铲子面积的函数，以此来确定最优铲子大小的思想，正是著名经济学家、管理学家弗雷德里克·泰勒（Frederick Taylor）提出的。在19世纪90年代，泰勒和其他一些人开创了科学管理的新纪元。在泰勒生活的那个时代，制造业决策，比如流水装配线的移动速度有多快，焊接强度要多高，工人的休息时间是多少，等等，都被建模为崎岖景观问题。进入20世纪后，许多伟大的实业家，例如福特汽车公司创始人亨利·福特（Henry Ford）、“石油大王”约翰·D. 洛克菲勒（John D. Rockefeller）和“钢铁大王”安德鲁·卡内基（Andrew Carnegie）都为这个现在用泰勒主义一词来概括的运动做出了卓越的贡献。

从制造个性化的、“只此一家别无分店”产品的工匠生产，转为大规模制造，是一个重大的变革。在大规模生产中，制造流程要分解为多个部分，每个部分都要经过优化，然后变成常规操作。这样一来，效率得到了大幅提高。但是在许多人看来，这个过程也是劳动的非人化过程。这种分歧就是一个提示：我们需要多个模型。任何单一模型都是对世界的过分简化，只能突出其中的某些维度。科学管理模型侧重于流程效率。这种偏向导致了批评。以产出效率为准则做出的决策，会导致其他目标遭到忽视，例如工人的快乐和福祉。

从表面上看，景观模型似乎只是一个相当浅显的想法：将适合度、效率或价值作为特征或性状的函数绘制在图上，然后爬上山顶，找到那个特征或性状的最优值。而且，把解决问题想象为“爬山”，似乎也不过是一个简单的比喻。这当然都是有效的批评。但是，如果构建了正式的景观模型，我们将能够得到一些非凡的结论。

崎岖景观

当我们同时考虑多个属性并且允许一个属性的贡献与其他属性的贡献相互作用时，就会得到一个崎岖景观，也就是具有多个山峰的景观。考虑一个设计沙发的问题，我们必须决定坐垫的厚度和扶手的宽度。我们用沙发在市场上的预期销售额来代表设计的价值，而沙发的销售额与设计的美感相关。如果沙发有厚厚的垫子，那么较宽阔的扶手可能会使沙发更具美感。如果沙发的垫子很薄，那么扶手窄一点会更好。作为扶手宽度和坐垫厚度的函数，预期销售的二维图将具有两个山峰。一个山峰对应于窄扶手、薄垫子的沙发设计；另一个山峰则对应于宽扶手、厚垫子的沙发设计。

变量之间的相互依赖效应，使得景观出现了崎岖的特点。这种崎岖性有好几个重要含义。首先，在崎岖景观中寻找到最高点时所用的不同方法，可能会以找到不同的山峰而告终。如果从不同的起点出发，也可能会找到不同的山峰。因此，崎岖性导致了对初始条件的敏感性和路径依赖的可能性。而这些都意味着，景观的崎岖性有助于结果的多样性。崎岖性也意味着出现次优结果的可能性，在崎岖景观中，次优结果表现为局部高峰。

图28-2显示了一个有5个山峰的崎岖景观。在这些山峰中，有4个是局部高峰，它们只是比与它们相邻的点的值高一些，只有一个是全局高峰，即具有最高值的点。要理解搜索是怎样止步于依赖初始搜索点的局部高峰的，可以想象从一个点开始往山峰爬的过程。这种过程被称为梯度启发式（gradient heuristic）或爬山算法（hill-climbing algorithm）。在崎岖景观中，梯度启发式技术会“卡”在局部高峰上。

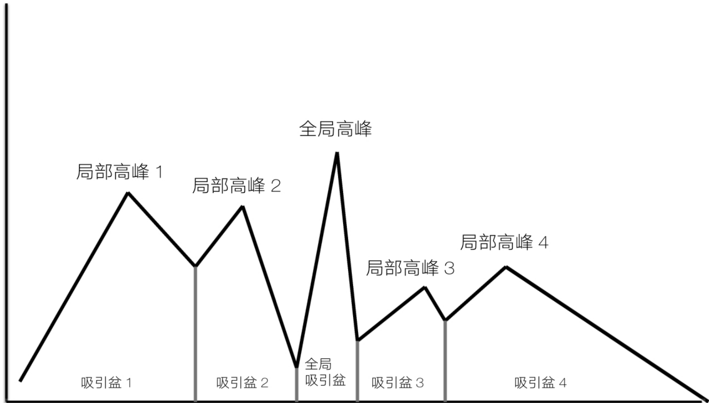


图28-2　一个有5个山峰的崎岖景观

如果起点位于最左侧，那么梯度启发式将定位于局部高峰1上，但它不是最优的。如果梯度启发式从图28-2中标识为“吸引盆2”的区域开始，那么它将定位于局部高峰2上。其他每个山峰，包括全局高峰，都有这样一个区域：如果梯度启发式从那个区域开始，就会找到那个局部高峰。这些区域被称为吸引盆（basin of attraction），如图28-2所示。从图中可以看出，全局高峰吸引盆的面积是最小的。如果我们随机选择一个起点并应用梯度启发式，那么全局高峰恰恰是最不容易被找到的那一个山峰。

吸引盆取决于启发式。如果我们使用了不同的启发式，就可能得到不同的吸引盆。例如，我们也可以不用梯度启发式，转而使用一个名为“一直向右走”的启发式。这个启发式一直向右侧搜索，直到找到一个局部高峰为止。对于这个例子，这两种启发式具有相同的局部高峰但却会产生不同的吸引盆，只要比较一下图28-3和图28-2，就可以看出这一点。

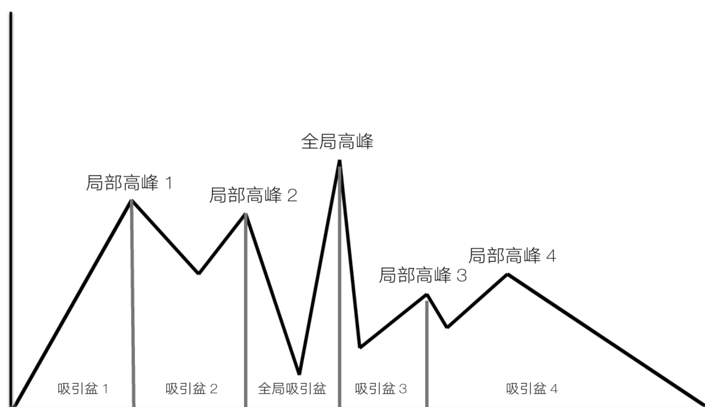


图28-3　由“一直向右走”启发式产生的吸引盆

要在崎岖景观中找到最优或接近最佳的高峰，需要多样性和复杂性的方法。多样性的价值是不言而喻的，如果不同的启发式能够找到不同的峰值，那么对一个问题应用多个不同的启发式就能够产生多个不同的局部峰值，然后就可以从这些局部峰值中选择一个最优的。[1]或者换一个思路，如果从不同的起点应用同一个启发式，那么也可以得到相同的结果：先找到若干个不同的局部最优点，然后选择其中最优的。

另外还应该注意到，景观的崎岖度（以山峰的数量来衡量）与问题

的难度相关。当然，有的问题很难解决，但却不一定有崎岖的景观。在一大块玉米地里找一枚金币这个问题可以用一个平坦的景观来表示，只是在硬币的位置有一个单峰。它的景观不是崎岖的，但是要找到硬币确实很难。

NK模型

现在描述NK模型。利用这个模型，我们可以对属性的相互作用与崎岖性之间的联系加以形式化。[2]这个模型将对象，或我们这里所称的备选方案，表示为一个长度为N的二进制字符串，这就是NK模型中“N”的含义。至于“K”，则指与该字符串的每一位交互以确定这一位的值的其他位的数量。如果K等于零，那么价值函数就是线性的。如果K等于N-1，则所有的位都相互交互，每个字符串的值都是随机的。因此，我们可以考虑增大K，将景观的崎岖度调整到富士山景观与随机景观之间的某个适当位置。

NK模型

一个对象由N位二进制符号组成，$s \in \{0，1\}^N$。

这个对象的价值表示为$V（s）= V_{k1}（s_1，\{s_{1k}\}）+ V_{k2}（s_2，\{s_2k\}）+ \ldots + V_{k1}（s_1，\{s_{2k}\}）$，其中，$\{s_{ik}\}$等于一个从原字符串中除了第$i$位之外随机选择出来的有$k$位的子字符串的集合，而且$V_{k1}（s_1，\{s_{1k}\}）$是从区间［0，1］中抽取出来的一个随机数。

$K = 0$：得到的是一个关于位的线性函数。

$K = N$-1：任何位的变化都使每个位产生新的随机贡献。

NK模型作为一个框架，为探索想法和提出问题创造了一个非常好的空间。我们想要问，局部最优值的数量是怎样取决于交互项的数量的。我们还可以问，全局最优值的大小又是如何取决于交互项的数量的。不过眼下，还不是回答这两个问题的适当时机，因为还没有定义好怎样去搜索可能性空间，这就是我们正在使用的启发式，局部最优集及其吸引盆的集合取决于搜索时所用的启发式。

在下文中，我们主要依赖单翻转算法（single flip algorithm）来进行搜索。这个算法按顺序选择每一个属性并切换该属性的状态。如果改

变那个属性会产生一个更高的价值，就保留这个切换。否则，那个属性就退回到它原来的状态。之所以选择这个算法，有两方面的动机。首先，我们可以把它解释为描述基因突变的一个粗略模型，让好的基因变异逐步接管整个种群，并让坏的基因变异消失。其次，它也是在我们这个空间中表征爬山算法最自然的一个形式。

作为例子，我们先来求解$N=20$且$K=0$时的NK模型。当$K=0$时，每个属性对总价值的贡献独立于其他属性，这时单翻转算法可以识别出每个属性更好的状态和全局最优值。因此，$K=0$意味着不存在相互作用，这种情况对应于富士山景观。每个状态的值均匀分布在区间［0，1］中。不难证明，从区间［0，1］上的均匀分布中随机抽取出来的两个值中较高的那一个期望值为2/3。只要我们对这20个属性的贡献求平均值，就不难推出全局最优值的期望值也是2/3。

而在另一个极端，当$N=K$-1时，每一个属性都与其他每个属性相互作用。在这种情况下，切换任何一个属性的状态，其他每个属性的贡献都会改变。那将是从区间［0，1］中均匀抽取出来的一个新随机数。同时对象的值则将是这20个新随机数的总和（每个属性一个）。这就意味着，属性的每一次翻转，都会导致整个对象产生一个新值，而且它与之前的值不相关。因此，景观将会非常崎岖，每一点都可能隆起、每一点都可能下陷。

在上面这两个结果的基础上，我们可以推导出局部峰值的期望数量。如果我们从任何一个备选方案开始，单翻转算法将会对该备选方案与N个备选方案中的每一个进行比较。例如，假设我们从所有位都取零值的备选方案开始，那么单翻转算法将评估N个备选方案，每一个备选方案都恰好有一位的取值为1。

初始备选方案：00000000000000000000

属性1切换后的备选方案：10000000000000000000

属性2切换后的备选方案：01000000000000000000

……

属性20切换后的备选方案：00000000000000000001

要成为一个局部高峰，其值必须高于这N个备选值中的每一个。原来的备选方案拥有最高值的概率等于$1/N$。因此，局部峰值的数量近似

等于可能的替代方案数量2^N除以N的商。很容易计算，对于$N=20$，大约有5万个局部峰值。由于具有如此之多的局部最优，只靠单翻转算法很难准确定位全局峰值。

因此，在这里重要的不是局部最优值的数量，而是它们的值有多大。因此还需要将这些最优值的期望均值与全局最优值的期望值进行比较。通过这种比较，我们可以确定单翻转算法的性能如何。而在计算这些值的时候，我们可以利用中心极限定理。在这个例子中，很容易证明局部最优值的期望值大约等于0.6，而全局最优值的期望值则略大于0.75。[3]将这些值与$K=0$时的全局最优值比较一下，就可以发现崎岖景观的局部峰值低于富士山景观的峰值，但是崎岖景观的全局峰值则比富士山景观的峰值更高。

由此自然而然地引出了这样一个问题：在这两个极端之间，也就是当我们将交互作用的属性数量K从零增加到N-1的过程中，发生了什么事情？答案是，我们可以观察两种效果。交互作用的属性数量的增加，在产生了更高的全局峰值的同时，也产生了更多的（更低的）局部峰值。假设我们在搜索时使用了单翻转算法，那么对这个模型的计算表明，对于比较小的K值，互动的好处（得到了更高的全局峰值）超过了互动的坏处（出现了更多的局部峰值）。因此，在开始阶段，局部峰值的期望值会随着K的增加而提高。同时，局部峰值数量的增加则意味着它们的平均值将减小。因此，如果决定使用单翻转算法，人们一般更喜欢用相对较小的K值，比如3或4。但是，我们为什么要把自己局限在这种只切换单个属性的简单启发式上呢？经由变异而完成的进化也许适用这个启发式，但是我们却不必受它所限。我们可以切换两个属性甚至三个属性的状态。更复杂的算法将会减少局部最优值的数量。

崎岖性与舞动的景观

NK模型的一个重要含义是，我们需要适度的相互依赖性，因为这种互动能够产生更高的峰值。多模型思维则要求我们跳出模型的特定假设，并仔细思考是什么原理驱动了这些结果。我们不难发现，背后的逻辑由两部分组成。第一个组成部分源于组合学：两个元素的组合数会随对数的平方而增加，而三个元素的组合数则会随三元组数量的立方而增加。因此，这种相互依赖效应有可能会创造出更多有益的互动。

第二个组成部分则源于我们只需要保持更好的组合这个事实。想象

一下，假设我们准备利用4种食材来制作一份食物。有4种食材意味着：如果选用其中的两种，那么有6种可能的组合。假设我们所用的是以下4种食材：泡菜，香蕉，鸡肉，焦糖。由此产生的6种配对是：香蕉和泡菜、泡菜和鸡肉、焦糖和泡菜、香蕉和鸡肉、焦糖和香蕉，以及焦糖和鸡肉。你可能会觉得只有一种搭配对你有吸引力，那么你就会选择那种搭配。假设我们喜欢的是焦糖和香蕉，就会忽略其余配对。[4]

类似的逻辑也适用于进化系统。能够产生正面相互作用的表型组合，例如坚硬的外壳与强健的短腿，会在种群中保存下来。适者生存法则与产生负面相互作用的组合相反。因此，我们现在看不到有如下这种组合的动物：跑得很慢的脚＋味道鲜美的皮肉＋外表鲜艳的外表。它们也许曾经存在过，但是早就被捕获并被吃光了。

我们在搜索模型中也会遇到类似的情况。当拥有非常多的可能性时，我们更喜欢变化。同样的逻辑在这里也是适用的：组合（两个元素、三个元素）会产生丰富的可能性。而且我们还希望，这些可能性的价值有很大的变化范围。然后，我们更有可能发现其中一个具有非常高的价值。由于相互作用效应会增加变异，因此总体上说，它们是有利的，但只在一定程度上有利。正如我们在上面已经阐述过的，太多的变化会使景观随机化。在理想情况下，我们会有适度的互动。有的学者认为，如果相互作用的数量和大小可以演化或适应，那么系统应该会自然而然地演变为具有高峰值的崎岖景观。如果真的是那样，那就表明系统倾向于向复杂性而不是均衡或随机性演化。[5]当然，什么时候能够达到这个结果、是不是真能达到这个结果，本身也是一个可以通过模型来探索的有趣问题。

最后要强调的一点是，我们一直将景观视为固定的。但是在生态和社会系统中，物种或企业要面对的景观还取决于他人的行为和属性。任何一个物种的适应，或任何一个企业策略的改变，都会改变和重组它们的竞争对手的适合度景观。

现在，我们可以将空间竞争模型和享受竞争模型重新解释为舞动的景观上的运动模型。这种运动可能会导致均衡，每个博弈参与者都站在局部或全局山峰上。或者，在舞动的景观上的竞争，也可能导致复杂的行动模式和结果。只要粗略地观察一下生态系统、政治领域和经济社会，就会明白后一种情况更容易出现。

我们之所以会观察到如此多的复杂性，一个很重要的原因可能是，我们这个世界在很大程度上是由自适应的、有目的的行为者组成的，它

们有能力操纵舞动的景观。为了理解这种复杂性，我们需要多模型思维。

我们可以对知识授予专利权吗

我们今天的幸福源于数百年来的知识积累。知识体现在所有方面：物理定律、内燃机、复式记账法、传染性细菌致病论、X射线和HTML等。知识通常是一种公共物品，永远是非竞争性的，不过知识既可能是、也可能不是排他性的。要想排他，必须有方法验证，当知识已经呈现为特定形式的人工制品时，要验证是比较容易的。例如，要想验证某个人是不是使用了某种算法或技术来解决了某个问题，那往往是不可能的；但是，要想验证某人是不是在软件程序中嵌入了某种算法，就是可以做到的。

当知识的排他性可以保证时，我们就会面临一个选择。我们可以像对待道路和国防那样去对待知识，并通过向民众征税来生产知识。政府可以向那些思考者支付报酬或者直接补贴他们，或者通过支持大学、研究机构来间接地加以支持。政府还允许人们获得专利权。专利制度之所以能够鼓励知识生产，关键就在于它为专利所有权人创造了一定期限的使用知识的独占权、并允许他们向其他使用专利的人收费。在美国和欧洲，专利权的期限为自申请提交之日起20年。[6]专利倡导者认为，如果任何人都可以免费使用别人发明的东西，那么私人（个人或机构）就不会有什么动力去开发更好的捕鼠器、计算机算法或音响系统了。他们强调，专利制度能够克服知识生产中固有的激励问题。

但是，经济学家米歇尔·博尔德林（Michele Boldrin）和戴维·莱文利用多模型思维，提出了一个有力的反对专利制度的理由。[7]在他们给出的允许思想（创意）组合的模型中，引入专利权会限制不同思想的组合，从而阻碍创新。如果一家公司获得了触摸屏技术专利，那么就很可能会减少其他企业设计采用这种技术的新产品的动力。如果没有专利保护，就会有更多的产品采用这种技术。也就是说，创新将会增加。

专利制度的支持者则反过来指出，就算专利制度真的会阻碍创新（那将很糟糕），但是如果没有专利保护，那么投资的减少幅度将会大得多。博尔德林和莱文基于我们在本书中讨论过的扩散模型反驳了这种说法。利用新知识设计生产的有用产品会迅速通过消费

者传播开来。收音机、电视和谷歌搜索引擎都是如此。这会创造出一种先发优势，创新者仍然会受益，尽管获利程度与专利保护下有所不同。

博尔德林和莱文还对某项发明应该在多大程度上归功于发明者提出了疑问。如果重大突破都是某个孤独的天才在密室中做出的，而且如果没有动力，大多数新思想都不会出现，那么专利制度就是有理由的。但是，崎岖景观模型表明，大多数困难的问题都有很多种可行的解决方案。新发明，特别是那些结合了现有思想和技术的发明，例如汽车、电话和在线拍卖，也许是本来就会发生的“自然事件”，而不是某个天才人物行为的结果。如果各种想法和创意都能够在思考者的群体中自由流动，那么很多人可能早就实现了这些创新。从历史上看，许多重大发明（发现）都有一种引人注目的同时性，例如，微积分是由艾萨克·牛顿和戈特弗里德·莱布尼茨发明的、电话是由亚历山大·格雷厄姆·贝尔和伊莱莎·格雷发明的，以及进化的自然选择理论是由查尔斯·达尔文和阿尔弗雷德·拉塞尔·华莱士发现的。

总而言之，多模型思维能够呈现专利制度的优点和缺点。这些模型提供的更深入、更周详的结论支持一种更加灵活的专利制度。也许，对于其中一些想法，那些许多人都能够发现的想法，以及可以与许多其他想法重新组合的想法，我们应该采取与今天的专利制度不同的专利制度，例如授予更短的保护期限、更宽松的使用条件。甚至，有些想法根本不应该被授予专利。

结语　像芒格一样智慧地思考——多模型思维的实际应用

> 一切都是复杂的；如果不是这样，那么生活、诗歌以及所有一切，都只会成为烦恼和负担。
>
> 华莱士·史蒂文斯

这是全书的最后一部分内容。在这一部分中，我们用多模型思维分析两个重要的政策问题：阿片类药物滥用和经济不平等。我们将阐明，如果同时运用多个模型，不但可以更好地分析这些问题，而且可以帮助我们理解为什么它们如此难以解决。我们还将会看到，特别是在阿片类药物滥用这个问题上，专家们如何利用多个模型在危机真的发生之前就预测到它。但是在这里，我们并不想说，利用模型就可以避免灾难的发生，那过于夸大其词了。我们这里对阿片类药物滥用问题的分析其实并不深刻，事实上，我们只是试图就如何利用多模型思维来考虑政策制定和政策实施问题给出一个粗糙的模板。我们没有收集过数据，也没有校准模型。相反，只是定性地应用模型得到一些见解。

然而另一方面，我们对收入不平等的分析则包括了更多细节，而且与学术文献紧密地结合在一起。它代表了多模型思维的另一个极端，对各个模型都深入地进行探讨。无论是对阿片类药物滥用问题，还是对收入不平等问题，利用多个模型来进行思考都会使我们变得更有知识、更聪明。但从定义上讲，复杂的系统是很难预测和理解的。我们肯定会犯错误，但是可以从这些错误中吸取教训，变得更加明智。

多模型思维与阿片类药物滥用

在美国，阿片类药物滥用情况有多严重？只要举出一个数字就够了：2016年，医生开出了超过2亿多张阿片类药物处方，这相当于差不多每个人一个处方。在那一年，美国有超过10万人死于与阿片类药物有关的服药过量。而有阿片类药物滥用问题的人则超过了1　000万人，其中有200万多人已经被归类为阿片类药物使用障碍患者。

医生之所以开出了如此之多的阿片类药物处方，主要原因当然是它

们有疗效，阿片类药物可以减轻疼痛。数以千万计的美国人都需要止疼，从而对这类药物产生了巨大的需求。但是没有人料到，这类药物会被滥用到这个程度。为了解释阿片类药物的滥用，我们采用了多模型思维方法。关于这个危机产生的原因，有四个模型都给出了一些重要的直觉性结果。

第一个模型是多臂老虎机问题，它解释了为什么阿片类药物会被批准使用。在申请药物上市时，制药公司要进行临床试验，以证明药物有显著的疗效并没有有害的副作用。我们可以将药物临床试验建模为一个多臂老虎机问题，其中一只手臂对应新处方药，另一只手臂则对应安慰剂或现有药物。

阿片类药物批准模型——多臂老虎机问题

为了证明阿片类药物的疗效，制药公司要进行药物与安慰剂的比对试验。在临床试验中，患者随机分为两组，一组服用阿片类药物，另一组服用安慰剂。我们可以把阿片类药物建模为一个“双臂老虎机”问题模型中的一只手臂、而把安慰剂建模为另一只手臂。在试验结束时，每个试验都归类为“成功”或“失败”。临床试验发现，接受阿片类药物治疗的患者的疼痛（在统计学的意义上）显著减轻了。对接受过髋关节置换术、牙科手术和癌症治疗的患者进行的药物试验都表明，阿片类药物的效果显著优于安慰剂。

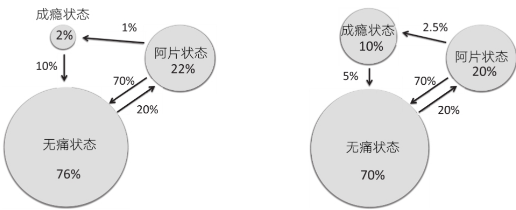
对于任何一种药物，成瘾的可能性都是一个非常值得关注的问题。只有当临床试验的结果可以证明，药物成瘾的患者比例极小（不到1%）时，这种药物才有可能获得批准。然而，这种试验并没有考虑过医生会给患者开出“大处方”的可能性，在某些情况下，医生一次开出的药就足够服用一个月。一个人服用阿片类药物的时间越长，成瘾的可能性就越大。数据表明，服药期较长的患者的成瘾率将会超过2.5%。下面的马尔可夫模型表明，成瘾率从1%提高到2.5%，就可以使阿片类药物成瘾者的均衡人数增加5倍。

这些转移概率仅用数据进行了不严格的校准。因为我们在这里用这个模型只是为了得出一个直观结论：相对较小的成瘾率是如何导致了大量成瘾者的。对这个模型的数值实验表明，只要我们稍稍降低成瘾者戒瘾成功的概率并提高从无痛苦状态转变为阿片状态的概率，那么阿片类药物成瘾者的比例会急剧增加。例如，如果我们在第二个模型中将从成

瘾状态转变到无痛苦的转移概率降低为1%，那么成瘾者的比例就会增加到35%。这种模型思维的含义是非常明显的，在现实世界中也开始得到了落实：有的医疗保健服务提供者，现在已经对医生可以开出阿片类药物的数量进行了限制。

成瘾模型——马尔可夫模型

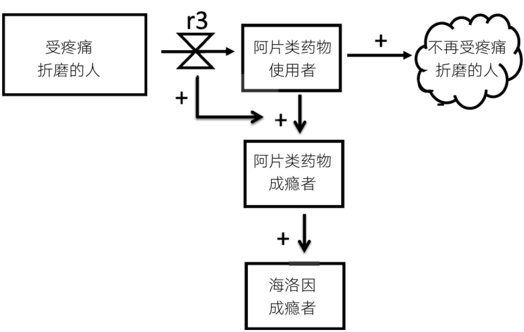
为了计算出成瘾的概率，我们创建了一个三态马尔可夫模型。这三个状态分别表示不受疼痛折磨的人（无痛状态）、使用阿片类药物的人（阿片状态）和成瘾者（成瘾状态）。我们要估计这三种状态之间的转移概率（用下图中的箭头表示。左侧的模型假设使用阿片类药物的人当中，会有1%的人成瘾，10%的成瘾者会恢复为无痛状态，并假设20%的处于无痛状态的人会成为阿片类药物使用者。在均衡状态下，只有2.2%的人是成瘾者。在考虑了“大处方”的情况下，左边的模型假设2.5%使用阿片类药物的人会成瘾，同时只有5%的成瘾者会恢复为无痛状态，并假设20%处于无痛状态的人会成为阿片类药物使用者。现在，在均衡状态下，会有10%的人是成瘾者。[1]



我们的第三个模型是一个系统动力学模型。与马尔可夫模型类似，在这个模型中，我们也假设了有三种类型的人（三个状态）：受疼痛折磨的人、使用阿片类药物的人和不受疼痛折磨的人。然而，这个系统动力学模型并不能直接写出这些状态之间的转移概率，而是想象存在一个“流”：由受疼痛折磨的人，到阿片类药物的使用者，再到不受疼痛折磨的人。更精细的系统动力学模型还可以包括其他一些源（其他药物的提供者），并允许阿片类药物使用者和海洛因使用者之间互动。此外，更精细的模型可能包括其他类型的情况。患有焦虑和抑郁的人更容易成瘾这一事实也可能被纳入这个模型。[2]

海洛因成瘾之路——系统动力学模型

模型描述了受疼痛折磨的人群产生阿片类药物使用者和海洛因成瘾者的过程。阿片类药物使用者会变为无痛状态，也会变为成瘾状态。而成瘾者则可以进一步发展成为海洛因使用者。人们使用海洛因的一个原因是他们不能再服用阿片类药物了。因此，随着阿片类药物的流量的加大，海洛因使用者的数量也在增加。



可以用来分析成瘾问题的最后一个模型是社会网络模型，我们没有在这里给出它的正式形式。这个模型依靠社会网络解释为什么人均阿片类药物使用量的地图会出现类似于农村地区那样的聚类。根据我们对平方根规则的分析，人口规模越小，变差越大。回想一下前面举过的例子：表现最好的学校和表现最差的学校都是小学校。农村地区阿片类药物的使用更大，也可能是因为医生会给农村病人开出服药时间更久的大处方，因为他们距离药店更远。除了这些解释之外，阿片类药物的使用也出现了聚类现象，而不再像是随机发生的。

如果阿片类药物是通过药物贩卖的渠道扩散的，那么就可能会出现这种聚类。与销售二手旧家具不一样（在销售二手家具时，人们会刊登广告），阿片类药物是通过个人关系销售出去的。因此，这样一个社会网络模型将会从家庭和朋友的社交网络开始。模型可能会假设，人们只向亲密的朋友出售阿片类药物。如果真的是这样，那么阿片类药物滥用者将会出现局部聚类现象。[3]

用多模型法分析经济不平等问题

接下来我们讨论最后一个多模型应用。在这里，我们将深入研究经济不平等的各种原因。之所以要讨论这个问题，有三个原因。首先，不平等是我们这个时代最重要的政策问题之一。收入和财富与人类社会的繁荣和未来息息相关。高收入人群拥有更好的健康状况、更长的预期寿

命、更高的生活满意度和幸福感。处于收入分布底部的人无论是谋杀率、离婚率，还是精神性疾病患病率都更高，他们普遍更加焦虑。[4]不过，我们必须非常小心，不要将相关关系与因果关系混淆起来。这种相关关系的很大一部分可以通过更健康、更快乐的人赚的钱更多这个事实来解释。但是，几乎所有研究都表明收入与生活幸福之间存在联系。没有人喜欢自己穷困潦倒。其次，我们有各种各样的不平等模型，这些模型分别由经济学家、社会学家、政治学家，甚至物理学家和生物学家提出。再次，我们拥有丰富的关于收入和财富的国内数据和跨国数据。我们不仅拥有当前的数据，还拥有可以回溯数百年的时间序列数据。

我们要先总结一下有关收入分配的若干经验规律。首先，在任何时代、任何国家，收入分布都有一个很长的尾部——许多低收入者和一小部分高收入者。在历史上，收入分布曾经被校准为对数正态分布或帕累托分布。最近，颗粒度更细的数据表明，收入分布的尾部长于对数正态分布，但是又不完全符合幂律分布。财富分布也同样是偏斜的。

其次，在大多数发达国家，近几十年来收入和财富不平等状况（无论用什么指标来衡量）一直处于不断恶化的趋势中。目前，美国的收入和财富不平等程度已经接近了镀金时代。由于整个分布内部的变化很难辨别，因此按照惯例，我们描述了归属于分布的上尾部的收入份额的变化情况。图29-1显示了顶层0.1%所占的收入份额随时间流逝而演变的情况。该图表明最顶层千分之一的家庭所占的收入份额在20世纪50年代后稳步下降，直到20世纪80年前后一直稳定在不到4%的水平。但是现在，这个数字已经上升到了10%左右。

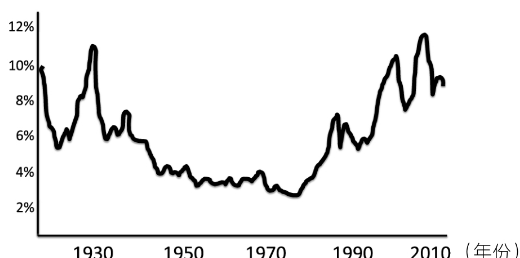


图29-1　最顶层0.1%的人的收入份额，1916—2010年。

资料来源：Piketty，2011。

再次，在全球范围内，生活在极端贫困中的人的数量也急剧下降了。我们应该看到，这些看似对立的趋势之间其实不存在逻辑矛盾。贫穷国家收入的快速增长，显著地减少了收入的跨国差异，但是并没有抵

消国内不平等程度的加剧。群体选择模型也产生了类似的效果。利他主义社区数量的增长，压倒了每个社区内部自私者上升的趋势。

不平等有很多原因，而且这些原因往往是相互交织的。经济力量、社会趋势、政治因素，以及发展历史，都会导致不平等。因此，正如经济学家史蒂文·杜鲁夫（Steven Durlauf）所指出的那样，我们不应该试图只用一个方程来解释所有不平等的水平或趋势，也不应该把所有政策建立在一个模型的基础上。[5]我们必须想得更加细致周到一些。使财富和收入集中在顶层1%或最顶层0.1%的过程，可能与将底层20%人困在贫穷陷阱中的因素无关。要深入了解收入分化的原因，需要采取多模型方法。

我们首先描述解释收入分配变化的模型。收入有以下几个来源：工资和薪金、营业收入、资本收入和资本收益。收入的这些组成部分股票的相对大小因收入水平而异。低收入人群的资本收益或资本收入很少。收入最高的那些人则从每个来源都可以获得很可观的收入。

我们的第一个模型扩展了柯布-道格拉斯（Cobb-Douglass）生产函数模型。在这个模型中，劳动包括两种类型：受过教育的人提供的劳动和未受过教育的人提供的劳动。支付给某种类型的劳动的工资取决于该类型的相对供给和技术。[6]这个模型可以解释近期基于供求关系不平等程度的上升。在20世纪50年代，制造业的增长增加了对未受过教育的工人的需求。与此同时，大学入学人数的大幅上升（这部分是因为美国《退伍军人安置法案》的实施），增加了受过教育的工人的供给。到了20世纪80年代，由于人们上大学的动力下降，减缓了大学毕业生数量的增长，再加上受教育程度较低的移民流入，增加了低技能工人的供给。与此同时，技术变革——自动化制造的兴起和向更加数字化经济的转型，增加了受过教育的工人的相对价值。他们的工资上涨反映了价值的这种变化。

按教育程度划分的平均收入时间序列数据与这个模型的拟合相当好。出于这个原因，许多经济学家依靠这个模型来提出政策建议。根据这个模型，他们主张增加所有人受教育的机会，因为这能抑制受过教育的工人工资的上涨趋势并减少不平等。这个模型很好地解释了总体趋势，但是它无法解释每个收入阶层内部的变差的扩大。

技术和人力资本模型——增长模型

产出取决于实物资本（K）、受过教育的人的劳动（S）和未受

过教育的人的劳动（U），具体生产函数如下：

$$产出 = AK^{\alpha}S^{\beta}U^{\gamma}$$

参数A、α、β和γ刻画了技术和三种投入要素相对价值。高技能工人和低技能工人的相对市场工资是：[7]

$$\text{Wage}_S = \left[\frac{U}{S}\right] \times \left[\frac{\beta}{\gamma}\right] \times \text{Wage}_U$$

（Wage_s是高技能工人的工资，Wage_s是低技能工人的工资）

不平等的原因：有利于受过教育的工人的技术变革会使β增大、使γ减少。这一点，再加上低技能工人供给的增加，会增加不平等。

不过，正反馈模型却可以。这个模型着重关注分布的尾部上的人群，特别是企业家群体。2011年，企业家在美国400名最富有的人中所占的比例达到了70%。[8]这个模型假设技术，特别是互联网和智能手机使人们之间的联系更加紧密，并使人们在更大程度上受他人的选择的影响。[9]一个想购买无线立体声扬声器的人，可以在线阅读评论，并从多个可选项中选出“最好的”。在过去，这个人也许只能在当地的立体声商店买，并只能选择唯一可选的某个型号。一个扭伤了膝盖的人，现在可以在网上搜索并了解他最喜欢的运动医生的身份。这种连通性会产生正反馈并导致更大的不平等。为了给这种受社会网络影响的经济选择建模，我们修改一下优先连接模型，以便将正反馈与人才联系起来，从而为受社会影响的经济选择建模。

虽然正反馈模型与时间序列数据的拟合程度不如前面那个强调技术导致增长的模型高，但是我们可以通过运行数值实验来了解反馈是如何导致不平等的。回想一下在第6章中描述过的音乐下载实验，也就是让随机分成两个组的大学生在两个不同的情境中下载音乐。在第一个情境中，被试们无法看到其他人在下载什么音乐。这种情境刻画了互联网出现前的世界。在第二个情境中，被试们可以看到每首歌的下载数字。在不能观察到“社交”信息的情境下，没有一首歌的下载次数超过了200次，只有一首歌的下载次数少于30次。然而，当人们可以看到下载次数时，有一首歌的下载次数超过了300次，同时一半以上的歌曲的下载次数少于30次。信息和社会影响放大了马太效应。富人变得更加富裕了，而穷人则变得相对更穷了。

对才华的正反馈——优先连接模型

存在N个生产者。开始时，每个生产者的销量均为零。第一个消费者随机选中了一个零销量的生产者，购买了产品，使该生产商的销量为正。随后，每个后续消费者都以概率p从销量为零的生产者那里购买、以概率（$1-p$）从具有销量为正的生产者那里购买。当从具有销量为正的生产者那里购买时，消费者选择生产者的概率与该生产者的当前销量成正比例。

不平等的原因：更多的联系增加了社会影响，创造了正反馈。

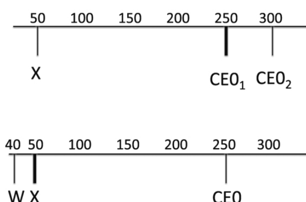
显然，我们可以将同样的逻辑应用于对经济问题的分析。[10]社交网络的正反馈效应导致不平等的可能性部分取决于人们所购买东西的性质。没有重量的商品，比如说可下载的电影、音乐和网络应用程序，以及某些技术，都很容易传播。点击一下图标是不能复制拖拉机、汽车和洗衣机的。因此，新的智能手机应用程序的销量可以几乎不需要付出任何资本支出就能够扩大，但是汽车却不能，即便是最畅销的汽车也不能。一个例子是，2015年5月，沃尔沃宣布，将在南卡罗来纳州生产S60轿车。这家新公司将于2015年9月破土动工，而第一批汽车将在2018年末才能下线。

下一个模型源于空间投票模型。我们用它来解释公司高管（如CEO）薪酬的上升，它不是由社会网络因素决定的。2012年，财富500强企业CEO的平均收入超过了1 000万美元，大约相当于当年工人平均工资的300倍。相比之下，在1966年，CEO的工资仅为工人平均工资的25倍左右。其他国家CEO的收入则要少得多。在日本，CEO的收入大约是普通员工的10倍。在加拿大和整个欧洲，CEO的工资大约是普通工人的20倍。

在大多数公司内，CEO的薪酬是由一个由董事会成员组成的薪酬委员会决定的，通常包括了工资、奖金和股票期权。这就是说，决定CEO薪酬的人往往是其他CEO。他们有很强的动机提高其他CEO的薪酬，进而提高自己的薪酬。我们可以使用一个空间模型来表示薪酬委员会的偏好。根据空间投票模型，工资将被设定为中间选民的偏好。CEO的薪酬跨国差异可以通过董事会和薪酬委员会的组成来解释。在德国，董事会包括工人成员，他们更倾向于减少CEO的工资。

CEO的薪酬——空间投票模型

CEO薪酬由薪酬委员会投票决定。在美国，薪酬委员会通常由现任和前任CEO（他们当然更喜欢高薪）以及薪酬专家（X）组成。而在其他国家，薪酬委员会的组成人员中还有工人（W），因而导致中间选民更偏好比较低的薪酬水平。



不平等的原因：CEO通过相互"俘获"来确定自己的薪酬。任何一位CEO的薪酬的增加，都会使所有CEO更加偏好更高的薪酬。

这个模型根据什么样的薪酬才是适当的来解释CEO薪酬的上升。在一点上，我们也可以回过头去考虑分析价值的多模型方法。薪酬委员会成员的理想价值可能是基于数据的，也可能在社会影响下形成的，或者也可能是基于对CEO未来价值的预测。

下一个收入不平等模型来自著名经济学家托马斯·皮凯蒂（Thomas Piketty）的畅销书《21世纪资本论》（*Capital in the Twenty-First Century*）。与其说这是一个正式的模型，还不如说它是一个观察结果：资本回报率总是会超过经济增长率。只要这个关系成立，那么高收入者从资本回报中获得的收入部分就会随着时间的推移而不断增加。如果在增长模型的基础上构建一个更精致的模型，应该不难证明资本回报率始终会超过整个经济体的增长率。从长远来看，经济增长率可能不到2%或3%，但是资本回报率则可能会高出一倍以上。

由此可以推出，在一个由赚取工资的工人和从租金中获得收入的资本家组成的经济体中，资本家的收入份额将会增加。更正式地说，资本增加的速度将取决于三个比率：消费率、税率和资本回报率。消费取决于资本存量水平。一个没有什么资本的人需要把自己收入的很大一部分消费掉，而拥有大量资本的人的消费只占收入的很小一部分。如下面的专栏所示，如果我们将消费率表示为一个常数除以资本水平，那么消费量将不会依赖于资本水平。较富裕的人将以较低的速度消费，而这会使他们的净资本更有可能增加。

资本租金模型（皮凯蒂）——**72**法则

经济由工人和资本家组成。工人的工资增长率为g，即经济增长率。资本家在时间t有财富W_t。资本的回报率为r（税后净额），且资本家的消费为一个不变的常数A。资本家的收入将比工人的增长更快，当且仅当：

$$r-\frac{A}{W_t}>g$$

不平等的原因：在市场经济中，资本回报率必定超过经济总增长率（即$r>g$）。拥有大量财富的资本家只将自己资本收入的一小部分用于消费，因此他们在总收入中的份额将随着时间的推移而不断提高。

为了说明资本回报率与经济增长率之间的差异是怎样导致不平等的，我们可以应用72法则。如果在最初的时候，工人的收入与资本家的收入相等，工资的年增长率为2%，而资本的年增长率为6%，那么在36年后，工资将增加一倍，但是源于资本的收入将增加8倍。而在72年内，资本家的收入将会达到工人收入的60倍。

皮凯蒂就是运用这样一个模型来解释收入和财富不平等的长期趋势的。用法国和英国过去三个世纪以来的数据对这个模型进行校准，结果非常好。这个模型还揭示了过去一个世纪在美国和欧洲各国的不平等的演变模式：两次世界大战摧毁了欧洲的资本存量，使那里的收入和资本分配较为平衡。这个模型能够很好地适合数据的一个原因是它略去了两个相互抵消的效应。由于将企业家排除在外，这个模型低估了不平等。而通过假设资本家的后代会明智地投资（尽管并非所有资本家的后代都能做到这一点），这个模型又夸大了不平等。因此，它在创造了一个新的富人阶层的同时，抹去了一个原来的富人阶层。这个模型，由于同时出现了这两个疏漏，因而比只有一个疏漏的模型更加准确。

这个模型的含义是，只要资本在增加，资本家就会从经济蛋糕中获得越来越大的份额。只要继续应用72法则，任何人都能看到资本家的收入将使工人的收入相形见绌。对于这种原因导致的收入不平等，似乎有一个非常简单的解决方案：向富人征收财富税。作为征税的替代方案，有的人可能会期待战争和革命，也就是以暴力的方式重新分配财富；或者，等待能够产生很多新资本家的技术突破的出现。

接下来要考虑的两个模型都优先考虑社会力量，而且两者都有很强的经验证据支持。第一个模型解释了选择性婚配，也就是所谓的门当户

对的婚姻选择所导致的不平等恶化。家庭的收入取决于夫妻双方的收入。如果一个低收入者与一个高收入者结婚，那么这种婚姻将有助于平衡收入分配。如果一个高收入者与另一个高收入者结婚，那么收入差距将会增大。虽然大多数人结婚时，都无法确知未来的人生伴侣的终身收入；但是，人们确实可以了解潜在结婚对象的受教育程度和身体健康状况，并且能够收到他们是否拥有雄心壮志的信号。有证据表明，男性（女性）受教育程度越高、收入越高（技术和人力资本模型），他们越有可能选择受教育程度较高的人生伴侣。

不平等的加剧是由以下因素造成的。第一，获得大学学位的女性越来越多。第二，相对收入会随着接受教育的程度的增加而提高。第三，受过良好教育的男性（女性）更喜欢受过良好教育的人生伴侣。因此，由两个受教育程度很高的人组成的家庭，更有可能拥有两个高收入者，从而导致家庭之间的收入不平等状况恶化。这里的逻辑似乎很严密。唯一的问题是这种效应的影响空间有多大。[11]

社会学家通常会将人们按教育水平分为五类：辍学、高中毕业、上过大学、有大学学位，以及研究生。然后，他们计算出每个教育水平的平均收入，并拟合关于每一对教育水平之间的婚姻数量的数据，从而粗略估计出选择性婚配的影响。

选择性婚配模型——分类模型与分类

每个人都有自己的教育水平：{1，2，3，4，5}。（其中1＝辍学，2＝高中毕业，3＝上过大学，4＝大学学位，5＝研究生）

令$P(m, j)$和$P(w, j)$分别表示男人和女人具有教育水平j的概率。收入(g, ℓ)表示性别为g、收入水平为ℓ的人的（估计）收入。一对夫妇组成的家庭的收入，包括了一个受教育程度为ℓ_M的男子和一名受教育程度为ℓ_w的妇女的收入。其家庭收入估计如下：[12]

$$收入(M, \ell_M) + 收入(W, \ell_w)$$

不平等的原因：受过良好教育的女性人数的增加、高教育水平的工人工资的增加，以及选择性婚配（人们喜欢与收入水平相同的异性结婚的倾向）导致了家庭之间收入不平等的增加。

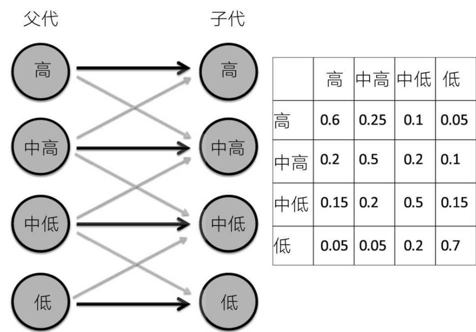
如果婚姻是随机的而不是非得门当户对不可的，收入不平等的程度就会轻得多。一项研究表明，如果婚姻是随机的，那么以基尼系数衡量

的收入不平等程度将会减少25%。[13]

下一个模型则使用马尔可夫模型分析不同收入类别之间的变动。这个模型按收入水平将人们（或家庭）分成四类：高、中高、中低、低。每个类别包括了分布的1/4。选定一个时间段，可能是1年、10年，也可能是一代人的时间，然后估计收入类别之间的转移概率，以刻画收入流动性。

代际收入（财富）动态变化——马尔可夫模型

将所有人口划分为4个人数相等的收入（或财富）类别。我们可以估计一个类别中的个体（或家庭）在一代的时间内流动到另一个类别中的转移概率（如下图所示）。更平等的转移概率对应更大的社会流动性。



	高	中高	中低	低
高	0.6	0.25	0.1	0.05
中高	0.2	0.5	0.2	0.1
中低	0.15	0.2	0.5	0.15
低	0.05	0.05	0.2	0.7

收入水平（收入等级）之间的转移概率

不平等的原因：社交技能、隐性知识、对风险和教育的态度以及遗产，减少了收入阶层之间的流动性。

如果代与代之间不存在黏性，那么高收入父母的孩子的收入属于4个收入阶层中的任何一个的概率都相同，即所有的转移概率都等于1/4。在完全不存在流动性的最极端情况下，转移概率将仅包括沿对角线的那些1。经验结果估计，表明现实介于这两个极端之间。

我们可以拿100个随机挑选出来的高收入或低收入水平家庭来进行仿真实验，计算出每一个后代收入的概率分布。如果使用上面的专栏中给出的概率，高收入者的子女有60%的机会成为高收入者，只有5%的机会成为低收入者，那么高收入者的孙子女成为高收入的可能性不足43%，成为低收入的可能性却会超过10%。[14]

这个收入动态模型也可以作为评估收入流动原因的基准模型。我们可能会使用一个线性模型来估计作为父母的财富、收入和能力水平的函数的子女收入（假设我们拥有相关的数据）。皮凯蒂的模型意味着父母的财富与子女的收入之间是正相关的。基于能力的模型则意味着父母的能力与子女的收入是正相关的，因为父母的能力和后代的能力之间存在某种相关性。

需要注意的是，要确定父母收入系数需要每个子女和每个父母的收入数据。只是过去几十年以来，我们才开始拥有比较完整的个人收入数据。幸运的是，在《太阳照常升起》（*The Sun Also Rises*）这本著名小说中，经济史学家格雷戈里·克拉克（Gregory Clark）发现了一种解决数据匮乏问题的新方法，这种方法利用了人的姓氏。克拉克计算了1888年所有名叫撒切尔的人的平均收入，并将之与1917年所有名叫撒切尔的人的平均收入进行了比较。30年的时间基本相当于一个人一辈子的工作时间。结果，克拉克发现姓氏的平均收入存在显著的相关性，这是表明收入缺乏流动性的一个证据。

有了这种模型，我们还可以识别代际转移中的种族差异。非洲裔美国人虽然也有进入收入分布顶端的，但是表现出了更低的财富持久性；而另一方面，他们在低端则表现出更大的持久性。这也就是说，富裕的非洲裔美国人不太可能拥有富裕的子女，而贫穷的非洲裔美国人却更有可能拥有贫穷的子女。[15]

我们要讨论的最后一个关于收入不平等的模型基于邻域效应，它是以杜鲁夫的持续不平等模型（persistent inequality model）为基础的，利用了人们按收入类别分离居住的经验规律性。这也就是说，高收入的人倾向于与其他高收入的人生活在同一个社区中，而低收入的人则与其他低收入的人住在一起。按收入类别分离居住会产生经济上、社会上和心理上的外部性，从而导致流动性下降。在这个模型中，个人收入取决于能力、教育支出和溢出效应。

教育属性还包含教育方面的公共支出。从经验上看，公共教育支出又与平均收入水平相关：高收入地区在公共教育方面的支出高于低收入地区，从而又可以为高收入社区的儿童带来更好的教育和更高的（未来）收入。

溢出可以解释为获取适当工具所需知识的社会传播。在这里，我们可以将杜鲁夫的模型与居住在高收入社区的人们如何获得关于何为适当工具的知识过程联系起来。还可以将模型与社会网络模型和“弱关系有大

力量”现象联系起来：生活在高收入社区的人们能够间接地与更多掌握了很有经济价值的信息的人联系起来。

我们还可以将溢出效应解释为一种社会传播行为，例如学习或工作所花费的小时数。如果收入包含了随机成分，那么低收入社区的人就会观察到，花费在自我完善、自我提高上的时间不会带来多少回报（这种观察结论是正确的）。与此相关，溢出还可能包括心理属性，比如对生活积极的或消极的态度、对社会的安全感，以及对自己的信念。

持续不平等（杜鲁夫）模型——谢林隔离模型＋局部多数模型

所有个人分别属于不同收入类别，并按不同收入类别分离居住。个人将自己的部分收入用于教育，从而产生正面的溢出效应，这种溢出效应随社区收入水平的上升而增强。生活在社区C中的孩子的未来收入取决于自己天生的能力、教育支出和溢出效应。教育支出和溢出效应的贡献取决于这个社区的收入水平，I_C。

个人收入$_C$＝F（能力、教育支出（I_C）、溢出效应（I_C））

不平等的原因：在低收入社区长大的儿童获得的教育机会较少、受益于经济溢出效应的可能性也更低。

在完整的模型中，杜鲁夫求解出了教育支出的均衡水平，并推导出了持续不平等产生的条件。这种持续的不平等源于他所称的“贫困陷阱”。生活在低收入社区的个人缺乏提高收入所必需的教育资源，当地也不存在可以令他受益的溢出水平，因此无论他们的能力水平如何，都很难走出贫困陷阱。杜鲁夫的模型有助于解释收入水平上的巨大种族差距，例如，非洲裔美国人生活在贫困社区的比例过高。杜鲁夫的模型解释了为什么他们更可能陷入低收入的人生轨迹无法自拔。

上面给出的这些模型突出了收入不平等的各种不同原因。对于收入分布的顶层，经验证据最支持的是那些以技术变革为基础的模型。[16] 20多年来，美国国税局追踪了收入最高的400个美国人的收入情况。最顶层这些人的收入主要来自新技术、大众零售和金融行业。这三个行业的共同特点是，它们都可以快速扩展。这种特别高的增长率可能源于像搜索引擎和社交网站这样的赢者通吃市场。但是，以技术变革为基础的模型不能解释收入分布底层的情况。关于收入流动性，它们也无法告诉我们什么信息。此外，它们也不能解释为什么美国CEO的薪酬远远超过了其他国家。

为了解释这些现象，我们需要引入其他模型，比如收入流动性模型、杜鲁夫的持续收入不平等模型和空间投票模型。只有在考虑了所有这些模型之后，我们对收入（财富）不平等才会有更加深刻且多方面的理解。我们看到，在不平等的产生和维持中，有许多不同的过程都起到了作用。我们还可以观察到，这些不同的过程之间存在重叠和交叉。当我们对不平等的复杂性和自我强化的因果关系有了更深刻的理解之后，我们就会对任何声称能够“快速”解决不平等问题的简单方案持怀疑态度。我们认为，减少不平等将需要在很多个方面共同努力。

让智慧入世

在本章中，我们讨论了如何将多个模型结合成一个整体来使用。利用这种方法，我们可以解释阿片类药物滥用和收入不平等的多种原因，并揭示任何一个解释框架的局限性。如果我们是制定政策的专家，就可以用这些模型中的某一个或某一些模型去拟合数据、衡量政策效果，还可以组织自然实验来指导政策选择。

我们还可以针对任何一种社会挑战来进行类似的分析，例如，扭转肥胖趋势，改善教学成绩，缓解气候变暖，管理水资源，甚至改善国际关系。

在每种情况下，即使添加一个新模型也可能产生巨大的后果。以预测金融崩溃为例。美国联邦储备委员会依靠传统的经济模型，使用通货膨胀、失业和存货等国民核算数据。但是这些数据存在滞后性，它们每周、每季度或每年发布一次。这些数据也来自调查，即整个经济的样本。

复杂性学者多恩·法默（J. Doyne Farmer）主张基于从网络上获取的实时数据创建第二类模型。这些新模型将依赖于更细粒度的实时数据，因此与传统的模型不同。法默认为，这样的模型可能比现有的模型要好得多，他可能是对的。然而，在预测和预防金融灾难方面，这些新模型并不需要更准确。考虑到新模型将使用不同的数据并依赖于不同的假设，它们将做出不同的预测。从多样性预测定理中我们知道，只要新模型的精度不差很多，当与现有模型相结合时，这些新模型将提高预测的准确性。用法默的话来说，政策制定者将会更有集体意识。

在做商业决策时，高管可能会从事类似的工作。高管可以应用多个模型来决定产品属性、产品发布时间，设计薪酬计划，构建供应链并预

测销售。因为这些操作都发生在一个复杂的系统中，所以任何一个模型都是错误的。多个模型将会带来更好的行动。

总而言之，当面临选择、预测或设计方面的挑战时，我们应该采取多模型方法。基于多模型思维的“谋定而后动”肯定要比仅仅基于冲动和直觉就行动更好。当然，这也就意味着，我们无法保证成功。即便有很多模型，我们也可能无法确定最相关的逻辑链。与要解决的问题有关的领域可能非常复杂，甚至利用许多个模型也可能仅能解释变差的一小部分。

在应用模型来辅助设计时，我们可能会发现自己无法构建出有用的模型抽象。在这种情况下，模型的简单性恰恰可能会成为它们失败的原因。面对复杂性，我们可能会发现模型无助于我们交流思想、做出准确预测和选择最优行动。我们的探索也可能是几乎没有价值的。这本书中讨论的模型的七大用途并不一定能提供顺利登顶的云梯。但是，即便是在这些情况下，我们也能受益。我们可以揭示出相互依赖性，能够理解为什么复杂的过程往往很难理解。

即便是在模型的帮助下，我们的推理能力也会受到限制，所以我们必须保持谦卑，也必须保持好奇心。我们必须继续构建新模型并改进现有模型。如果某个模型遗漏了世界的某些关键特征，例如社会影响、正反馈或认知偏差，那么我们就应该构建能够包含这些假设的其他模型。当采取这些方法后，就可以分辨出哪些属性在什么时候是重要的、重要程度有多高。所有模型都是错的，这当然是事实。但是，我们不会因这个事实而泄气；恰恰相反，它会成为我们通过多个模型来追求智慧的动力。

同样重要的是，我们还应该在这种努力中追求乐趣。虽然本书一直在强调一些务实的目标，比如成为更好的思考者、在工作中取得更好的成绩、做一个有知识有智慧的世界公民。但是，它同时也隐含着另一个重要的目标，那就是，揭示建模的乐趣。建模的实践可以成为一个非常美丽的邂逅。我们做出假设、制订规则，然后根据规则、运用逻辑。正是通过这种合乎逻辑的努力，我们才能提高自己、使自己变得聪明。我们要将这种智慧带入世界，并积极地利用它去改变世界。

译者后记

当今时代，大数据至热。但是许多人其实只是被大数据裹胁了。有了数据还远远不够，我们还要懂得如何让数据告诉我们信息，懂得如何从信息中总结出知识。说到底，我们需要形成自己直面世界的思维体系。斯科特·佩奇这本书为我们指出了实现这个目标的方向和路径：多模型思维。

佩奇是个跨学科天才，他在书中给出的模型，涉及各个学科领域。有些模型其实相当复杂，但是佩奇总能说得既简明又透彻。在《多样性红利》一书中，佩奇强调了群体智慧可以改善我们的决策；在这一本书中，他进一步落实了如何实现多样性红利，那就是，利用多个模型，让模型组成一个群体来帮助我们。这本书是一个完整的工具包。我相信，只要你愿意尝试成为模型思考者，就肯定能从本书中收获很多东西。

当然，也要防止另一个极端。“只要不断地拷打数据，数据总会说话”，类似的说法已经屡见不鲜。我们理当保持警惕。不要忘记佩奇写这本书的思想前提和知识基础。他是复杂性科学的大师级人物，已经到了驭繁如简的境界。模型就是帮助我们像大师那样去思考的工具。但是，我们切不可把我们要面对的世界简单化。不要以为在土里插个温度计，就可以测量整个地球的温度了。恰恰相反，模型思维的前提就是承认复杂性。作为本书核心的多模型范式，实际上就可以说是复杂性的另一种表述。

译书从来不易。我首先要感谢的是我的太太傅瑞蓉。本书与我以前的每一本译著一样，她的贡献至少要占到一半。感谢儿子贾岚晴带给我的动力和快乐。这本书献给他们。

感谢钟睒睒，有情怀、重情义、善创新、爱读书的他，既是好老板，也是好老师，让我在工作之余能够腾出时间来完成此书。

我还要感谢汪丁丁教授、叶航教授和罗卫东教授的教诲。我对复杂性科学的兴趣，源于他们。同时还要感谢何永勤、虞伟华、余仲望、鲍玮玮、傅晓燕、傅锐飞、陈叶烽、童乙伦、罗俊、邓昊力、陈姝、黄达强、李燕、李欢、丁玫、何志星等好友的帮助。

这本书，是我翻译的第二本佩奇的著作。佩奇的书，简洁而丰富，纯粹而复杂，因而总是有一种神奇的吸引力，我翻译起来也觉得特别得

心应手。这是非常难得的体验。感谢湛庐文化和简学给我这个机会。

书中错漏之处在所难免，敬请专家和读者批评指正。

贾拥民

于杭州耑谷阁

注释

01　做一个多模型思考者

1. 例如：凯西·欧奈尔（Cathy O'Neil）在2016年阐述了基于数据的简单模型如何忽略人口的某些部分以及本书第4章提到的自适应。

2. 请参见帕阿斯克和希勒分析森林工业的论文（Paarsch &Shearer, 1999）。在他们的模型中，木材加工厂主和林场主要向工人支付按树计件工资。关于植树造林的原始数据显示，计件工资越高，工人种植的树木越少；向某个种植树木的人支付的报酬越多，这个人种植的树木就越少。直接得出的推论与标准经济学原理背道而驰——经济学理论告诉我们，如果每种植一棵树都可以得到更多的报酬，那么种植树木的工人就应该更加努力地工作才是。

   不过，如果我们构建一个包含了小时工资率的市场模型，假设每小时的体力劳动报酬为20美元，那么数据和我们的直觉就都可以证明是正确的。从这个假设可以得出这样一个等式：

   20美元＝每小时种植的树木数量×种一棵树的美元报酬

   这个模型意味着，如果树木很难种植，那么工资将会很高。如果一个人可以在一小时内种植10棵树，那么种每棵树的工资就等于2美元。如果树很容易种植，那么工资就会很低。如果一个人可以在一小时内种植20棵树，那么种每棵树的工资将为1美元。因此，这个模型预测，计件工资率与种植的树木数量之间呈负相关关系。这样一来，数据中看似矛盾的东西通过这个模型的镜头来看，在逻辑上是解释得通的。这个模型通过种植条件的变化来解释相关性。事实上，这个模型预测的不仅仅是这种相关性，它还给出了一种特定的数学关系：计件工资率乘以种植的树木数量应该等于一个常数。我们可以对这个结果进行检验。

3. 关于模型胜过了人的证明，请参见：Dawes, 1979；Tetlock, 2005；Silver, 2012；Cohen, 2013。关于认知偏差，请参见：Kahneman, 2011。

4. Slaughter, 2017 & Ramo, 2016.

5. 研究表明，最具影响力的科学发现和专利是不成比例地从多个学科中汲取灵感的。有一项研究在分析3 500万篇论文后指出，从长远来看，跨学科论文的影响更大（Van Noorden, 2015）。思想的组合不一定会表现为模型的组合，但在许多情况下确实是如此；请参见：Jones、Uzzi&Wuchty, 2008; Wuchty、Jones&Uzzi, 2007。如果我们将专利解释为创新的证据，那么有两个独立的研究思路都可以将思想的多样性与成功联系起来。Shi、Adamic、Tseng、Clarkson（2009）的研究表明，跨类别的专利被引用得更多。Youn、Strumsky、Bettencourt、Lobo（2015）则证明，大多数专利都包括了多个子类别。跨学科研究也稳步发展到了这样一个临界点：平均而言，社会科学家引用其他学科的论文，多于他们所属学科的论文。

6. Box &Draper, 1987.

7. Page, 2010a.

8. 我这样说并不意味着可以将知识与模型等同起来。我的意思是，模型可以表示知识并提供一个清晰的形式来交流对有关知识的理解。“知识”一词的含义非常广泛，它还包括了嵌入在身体中的、隐性的技能，例如怎么打网球、说法语或讨价还价。我在本书中使用的是更狭义的知识定义。关于更宽泛的知识概念，请参见：Adler, 1970。

9. 只要注意到，从高空往下掉落的跳伞者在终点时速度可以达到320千米／小时，就可以得出这个近似结果。终端速度与质量的倒数成比例。假设跳伞者的质量比毛绒玩具猎豹大400倍。400的平方根等于20，因此，毛绒玩具猎豹的终端速度大体上等于320千米／小时除以20，即大约16千米／小时。

10. 他没有说错。弗雷斯诺市比冰岛还要大30%。埃里克·鲍尔和约瑟夫·李普曼（Joseph LiPuma）（2012）讨论过怎样把学术研究得出的教训应用到商业世界中。

11. Lo, 2012.

12. 这个故事的不同版本在美国哲学家与心理学家威廉·詹姆斯（William James）、斯蒂芬·霍金和法学家安东宁·斯卡利亚（Antonin Scalia）的著作中都可以找到。

02 模型的7大用途

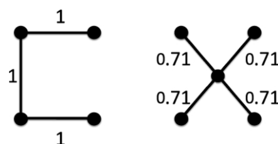
1. 关于对构建模型原因的更细致的分类剖析，请参见：Epstein，2008。查尔斯·拉夫和詹姆斯·马奇（1975）描述了模型的3个用途：解释经验现象、预测其他现象和新的现象，以及建造和设计系统。他们也都提倡利用模型进行探索。

2. 请参见：Harte, 1988。这种方法也借鉴了一篇阐述的关于模型在社会科学中的各种用途的观点的论文（Johnson, 2014）。这两种方法通常也称为伽利略和极简主义的理想化。请参见：Weisberg, 2007。关于类比的作用更进一步的讨论，请参见：Pollack，2014；Hofstadter&Sander，2013。侯世达在《表象与本质》一书中将类比称为思考之源和思维之火。另外也请参见：Schelling，1978、1987；他也详细地说明了模型的各种类别。丹尼尔·利特尔（Daniel Little）的博客“理解社会”（understanding society）也提供了相当不错的社会科学本体论的入门知识。

3. 请参见：Arrow，1963。如果我们对可能的个人排序施加限制，那么进行集体排序就是有可能的。比如，如果令每个人都保持相同的排序，那么集体的排序也会如此。但是一般而言，我们无法将个人排序映射为一致的集体排序。

4. 与我同一时代的杰出人士一定已经注意到了，我从《嚎叫》（*Howl*）中借用了“这是真的发生过的事情”的手法。请参见：Bickel、Hammel&O'Connell, 1975。

下图是一个添加一个节点会减少总边长的诸多可能例子之一。左侧的网络有4个点（分别作为正方形的一个角）。右侧的网络则多了位于中心的第5个节点。如果将正方形的边的长度假设为1，那么左侧图中各边的总长度为3，右侧图中各边的总长度则等于4×0.71，小于3。



当向录取率更高的院系提出入学申请的男性多于女性时，辛普森悖论就会出现。例如一所拥有医学院和兽医学院的大学。假设有900名男生申请入读医学院，其中480名（53%）被录取；同时有300名女生申请医学院，180名（60%）被录取。并假设有100名男生申请兽医学院，20名（20%）被录取；同时有300名女性申请兽医学

院，90人（30%）被录取。在每一个学院，女生的录取率都比较高，但是就整个学校总体而言，男生的录取率达到了50%（1000人申请，500人被录取）；女生的录取却只有45%（600人提出申请，270人被录取）。

作为“帕隆多悖论”的一个例子，假设有这样两个赌局：在第一个赌局中，一个人总是会输1美元；在第二个赌局中，如果轮数不能被3整除，那么输2美元，如果能够被3整除，就赢3美元。每一次投注都会形成一个期望损失，但是如果你只在赢的轮次上参加第二个赌局，而在其他轮次上则参加第一个赌局，那么你每三轮就肯定能够赢得1美元。

5. Kooti、Hodas&Lerman, 2014.

6. 假设每个人都可以赚得相同的收入，并以固定税率t缴纳税收。令c表示百分比减量、r表示收入的增量。当前的政府收入等于$I \times t$。减税政策实施后，政府的收入等于$I(1+r) \times t(1-c)$。只有当$I \times t < I(1+r)t(1-c)$时，政府收入才会增加。重新排列各项，我们有$r > c(1+r)$。

7. 基于市场的机制如何能为多维有效载荷问题提供一个更好的解，请参见：Ledyard、Porter&Wessen, 2000。

8. 物理学家尤金·维格纳在他的论文（Eugene Wigner, 1960）中使用过的一个形容词“好得不合情理”。他说，物理科学中使用数学模型的有效性“好得不合情理”。

9. Ziliak&McCloskey, 2008.他们讨论了社会科学模型解释变差的能力。

10. Porter&Smith, 2007.

11. Squicciarini&Voigtländer, 2015.Mokyr, 2002.

12. 请参见美国财政部网站。

13. 例如，在20世纪90年代中期，在美国俄亥俄州哥伦布市新开业的餐馆中，大约有60%后来都关门了。没有任何一家得到过政府的救助，当然政府也不应该救助它们。健康的市场经济本身必定包含了破产。请参见：Parsa et al., 2005。

14. 摘自国际货币基金组织的《2009年全球金融稳定报告》。连接的强度基于投资组合价值的相关性。但是这种相关性又是基于极端事件的，也就是这些机构表现得特别好或特别糟糕的日子。这个指标可

以刻画一家金融机构的破产扩散到另一家的可能性。事实上，业绩表现的相关性可能源于投资组合的相似性以及一家银行对另一家银行资产的持有。

15. Geithner, 2014.

16. 要了解旧金山湾的物理模型及其在政策中的用途，请参见：Weisberg, 2012。

17. Stone et al., 2014.

18. Dunne, 1999&Raby, 2001.

03 多模型思维

1. Levins, 1966.

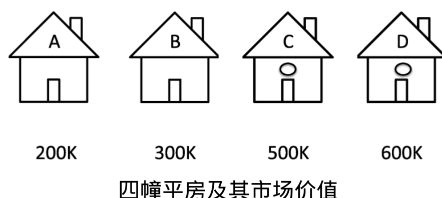
2. Page, 2007、2017.

3. 关于群体的智慧，请参见：Suroweicki, 2006；关于狐狸是如何胜过刺猬的，请参见：Tetlock, 2005；关于计算机科学的集成方法，请参见：Patel et al., 2011。

4. 卢红和佩奇（2009）证明，独立的模型需要一个唯一的分类集。这也就是说，在二元分类模型中只存在一种创建一组独立预测的方法。

5. 关于多样性预测定理的详细阐述，请参见我以前出版的几本书：《多样性红利》《多样性与复杂性》等。关于经济预测所用的数据，请参见：Mannes、Soll&Larrick, 2014。

6. 考虑下图中用A、B、C和D表示的四幢平房及其市场价值。根据这些平房是否包含了一个录音室由“门”上方的圆圈符号表示），创建两个类别。平房A和B不包含录音室，因此它们属于同一个类别；而平房C和D则包含了一个工作室，因此它们属于第二类。



四幢平房及其市场价值

首先计算出这些平房价格的总变差，即每幢房子的价值与均值之差

的平方和。这四幢平房的均值等于400 000美元，因此总变差等于100 000：总变差＝（200-400）^{2}＋（300-400）^{2}＋（500-400）^{2}＋（600-400）^{2}。

为了计算分类误差，假设我们知道每个类别内的真实均值：第一个类别为250 000美元，第二个类别为550 000美元（平房已转变为录音工作室）。但是这种分类将不同价值的房子混同在了一起。剩余的变差等于分类误差：分类误差A和B＝（200-250）^{2}＋（300-350）^{2}＝5 000、分类误差C和D＝（500-550）^{2}＋（600-550）^{2}＝5 000。总分类误差等于这两个数字的总和，即10 000。

为了计算出估值误差，假设模型预测平房A和B的价值均为300 000美元、同时预测平房C和D的价值均为600 000美元。估值误差等于每个类别的预测值与真实均值之间的差的平方和。估值误差A和B＝（300-250）^{2}＋（300-250）^{2}＝5 000，估值误差C和D＝（600-550）^{2}＋（600-550）^{2}＝5 000。于是总估值误差等于10 000。

总模型误差等于预测值与实际值之间的差的平方和：模型误差＝（200-300）^{2}＋（300-300）^{2}＋（500-600）^{2}＋（600-600）^{2}＝20 000。因此不难注意到，模型误差等于分类误差和估值误差之和。

7. 读者如果想了解二维社会互动的旋转玻璃模型，请参见：Brock&Durlauf, 2001。爱德华·格莱泽（Edward Glaeser）等人则给出了一个关于犯罪的一维模型。另外，德鲁·弗登伯格（Drew Fudenberg）和戴维·莱文（David Levine）也构建了一个关于大脑的经济模型。

8. 斯塔夫罗斯·尼阿科斯并不是第一个试图利用货船超大空间的人。1858年，因设计建造大西部铁路的传奇英国工程师伊桑巴德·金德姆·布鲁内尔（Isambard Kingdom Brunel）建成了一艘近220米长的大船，并命名为“大东方号”（SS Great Eastern）。事实证明，作为一艘货轮，大东方号是失败的。由于缺乏适当的水动力模型，导致大东方号整体设计不佳。这艘船只有在最慢的速度下才能够航行。幸运的是，它最终还是派上了用场——被用于铺设跨越大西洋的海底电缆。关于多模型思维在设计船舶时的作用，请参见：West, 2017。

9. 以磅和英寸为单位时，$身体质量指数=\frac{703\times体重}{身高的平方}$。

10. 勒布朗·詹姆斯（LeBron James）身高6英尺8英寸，体重大约250磅，身体质量指数为27.5。凯文·杜兰特（Devin Durant）身高6英尺9英寸，体重235磅，身体质量指数为25.2。2012年和2016年奥运会十项全能金牌得主阿什顿·伊顿（Aston Eaton）身高6英尺1英寸，体重185磅，身体质量指数为24.4。另一位著名运动员，2008年奥运会十项全能金牌得主布莱恩·克莱（Brian Clay）的身体质量指数则为25.8。

11. Flegal et al., 2012.

12. 假设这只小鼠3英寸长、1英寸高、1英寸宽，同时假设这只大象10英尺高、10英尺长、5英尺宽。大象的表面积为400平方英尺，也就是57 600平方英寸。大象的体积为500立方英尺，相当于864 000立方英寸。

13. 杰弗里·韦斯特（Geoffrey West, 2017）和他的同事们已经构建了一个更精细和更准确的模型，它预测新陈代谢的规模与质量的四分之三次幂成比例。而且在80多年之前，科学家就已经在数据中找到了与之相近的定律，即克雷伯定律（Kleiber's law）。

14. 发送的简历内容完全相同，但是申请人的名字改变了的实验表明，女性获得的报酬低于男性，得到的评价也低于男性；请参见：Moss-Racusina et al., 2012。

15. 一个男性员工成为CEO的概率等于他连续获得15次晋升的概率。概率比，也就是男性成为CEO的概率与女性成为CEO的概率之间的比值。按照我们的假设，男性每次升职的概率为50%，女性每次升职的概率为40%，于是概率为$(1.25)^{15}=28.4$。

16. Dyson, 2004.

17. Breiman, 1996.

04 对人类行为者建模

1. Haidt, 2006; Simler&Hanson, 2018.

2. 要证明这个结果，需要利用一些微积分知识。假设一个人的收入为

M，一单位食品的价格为1美元，一个单位的住房服务的价格等于P_H。那么我们可以将他的预算约束写成$M=F+P_H\times H$。而这当然就意味着$F=M-P_H\times$H。可以将其效用函数写成

$$U(H)=(M-P_H\times H)^{\frac{2}{3}}H^{\frac{1}{3}}$$

为了求这个函数的最大值，我们对H求导数并令它等于零：即

$$-\frac{2}{3}P_H\left(\frac{H}{M-P_H\times H}\right)^{\frac{1}{3}}+\frac{1}{3}\left(\frac{M-P_H\times H}{H}\right)^{\frac{2}{3}}=0$$

把第一项移到等式的右侧，交叉相乘，可以得到：$2P_H\times H=M-P_H\times H$。在预算约束的条件下，用$2P_H\times H$替换$M-P_H\times H$，可以得到$M=2P_H\times H+P_H\times H$，或者$M=3P_H\times H$。因此，他将收入的1/3用于住房。

3. 在美国，这是一个不错的近似值。

4. 这个定理的正式表达如下：$X=\{$A，B，C，…，$N\}$表示一个结果的有限集，并设定一个抽奖，令它成为这个结果集上的一个概率分布，即$L=$（pA，pB，…，pN）。如果对彩票的偏好（>）满足完备性、传递性、独立性和连续性，那么这个偏好可以表示为一个连续的有实数值的效用函数，可以给每个彩票赋予一个效用。完备性是指，任何两个彩票，L和M，都可以进行比较；传递性是指，如果$L>M$，并且$M>N$，那么$L>N$）；独立性是指，如果$L>M$，那么给定任意一个彩票N和任意一个概率$p>0$，都有$pL+$（$1-p$）$N>pM+$（$1-p$）N；连续性：如果$L>M$且$M>N$，那么存在概率p使得$pL+$（$1-p$）$N\sim M$。这个定理的正式证明有点繁杂，下面只给出证明的简要思路。假设存在最优结果B和最差结果W，并且假设B的效用等于1、W的效用等于零。给定任何其他结果A，根据连续性公理，存在概率p，使得人们在肯定能够得到A，与以p的概率得到B并以（1-p）的概率得到W，两者之间无差异。利用上面的符号，我们把它写成A$\sim p$B+（1-p）W。然后，我们再将A的效用指定为p。按此思路推导（这需要一点点数学知识），可以证明，一个人越喜欢某个结果（或彩票），p就越大。这样一来，我们也就魔术般地将排序转变成了数字。对于这个定理的完整的证明，请参见：Von Neumann&Morgenstern, 1953。

5. Rust, 1987.

6. Camerer, 2003.

7. Harstad&Selten, 2013.

8. 关于以理性选择模型为基准的理由，请参见：Myerson, 1999。

9. 对于这个领域的一个早期的综述，请参见：Camerer、Loewenstein&Prelec, 2005。

10. Kahneman, 2011.

11. 开放科学协作组织的原始论文（2015），导致了更多的复制结果的研究，它们得出相似的百分比。

12. 要想了解关于被试多样性的重要性，请参见：Medin、Bennis&Chandler, 2010。

13. Berg&Gigerenzer, 2010.他们充分表达了这种批评意见，认为它破坏了基于数学的心理模型。

14. Kahneman&Tversky, 1979.

15. 例如，治疗方案A——肯定能够拯救40%的患者，治疗方案B——有50%的机会拯救所有人。损失框架：治疗方案A'——肯定会有60%的患者死亡，治疗方案B'——有50%的机会没有任何人死亡，有50%的机会所有人都死亡。根据前景理论，大多数医生在收益框架下都会选择治疗方案A，而在损失框架下都会选择B'。

16. 关于双曲贴现的含义的早期研究，请参见：Thaler, 1981；Laibson, 1997。

17. 双曲贴现的公式可以写成如下更一般的形式：

$$H(r,t)=\frac{1}{(1+rt)^{\frac{\beta}{\alpha}}}。$$

18. Gigerenzer&Selten, 2002.

19. Gode&Sunder, 1993.

20. Gigerenzer&Selten, 2002.

21. 在弗农·史密斯（Vernon Smith）的诺贝尔奖获奖演说中，他指出："生态理性，要使用推理，也就是用理性的重建来考察个人的行为，但根据的是他们的经验和常识。人们遵循规则但不一定能清楚地将它们表达出来。但是，它们是可以被发现的。"请参见：Smith,

2002。

22. Arthur, 1994.

23. Lucas, 1976; Campbell, 1976.

24. 关于模型如何阐明可能发生的事情的进一步讨论，请参见：de Marchi, 2005。关于基于方程和规则的模型（行为者在类似的博弈中做出类似的行为），请参见：Gilboa&Schmiedler, 1995；Bednar&Page, 2007、2018。

05 正态分布

1. 给定一个数据集{x_i，…，x_N}，方差等于数据与均值之间距离的平方的平均值，可以写为：

$$\text{VAR}=\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$$

标准差则等于数据与均值之间距离的平方的平均值的平方根，即：

$$\text{s.d.}=\sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$$

2. 这里有好几个条件，而且只要满足其中任何一个条件就足够了。林德伯格条件（Lindeberg condition）是比较常见的一种，它要求，随着变量的数量的无限增加，来自任何一个变量的总变差的比例将收敛为零。

3. Lango et al., 2010.

4. 对于一般情况，假设独立性，我们可以得到如下表达式：

$$\bar{\mu}=\frac{\sum_{i=1}^{N}\mu_i}{N}$$

$$\bar{\sigma}=\sqrt{\frac{\sum_{i=1}^{N}\sigma_i^2}{N^2}}$$

用σ代替σi，就可以得到$\bar{\sigma}=\sqrt{\frac{N\sigma^2}{N^2}}=\frac{\sigma}{\sqrt{N}}$。

5. Wainer, 2009.

6. 两个标准差的阈值（5%显著性）是一个惯例。尽管经常受到批评，但它仍然是社会科学家最常使用的。当然，6%显著性水平的大相关系数，可能比4.9%显著性水平的小相关系数更值得注意。进一步的讨论，请参见：Ziliak&McCloskey, 2008。

7. Gawande, 2009.

8. 随机变量的乘积分布之所以称为对数正态分布，是因为分布的对数将会是正态分布的。个中的原理，我们可以简述如下。首先，我们写出一些数字的乘积，$y=x_1\times x_2\times x_3\ldots\ldots x_n$，并将它写为如下以10为底的幂的乘积的形式：

$$10^{log_{10}(y)}=10^{log_{10}(x_1)}\times 10^{log_{10}(x_2)}\times 10^{log_{10}(x_3)}\ldots 10^{log_{10}(x_n)}=10^{log_{10}(x_1)+log_{10}(x_2)+log_{10}(x_3)+\cdots+log_{10}(x_n)}$$

然后，我们对两边同时取以10为底的对数，得到以下的等式：

$$log_{10}(y)=log_{10}(x_1)+log_{10}(x_2)+log_{10}(x_3)+\ldots+log_{10}(x_n)$$

因此，变量y的对数可以写为随机变量的对数之和。这些随机变量的对数也是随机变量，只要它们的方差满足中心极限定理的条件，它们的和$log_{10}(y)$就是正态分布的。

9. Limpert、Stahel&Abbt, 2001.

10. 这个思想最早可以追溯到罗伯特·吉布拉特（Robert Gibrat）1931年的研究。

06 幂律分布

1. 关于这场大洪水的影响和文化意义，请参见：Parrish, 2017。

2. 在这个例子中，指数假设为2。这个例子借用自：Clauset、Young&Gleditsch, 2007。

3. 关于本章所述模型的技术性细节以及其他关于幂律的例子，请参见：Newman, 2005。

4. Newman, 2005; Piantadosi, 2014.

5. 常数C使所有结果的总概率等于1。根据这个定义，幂律分布满足标度不变性（scale invariance），如果改变了用来衡量结果的单位，那么分布的形状仍然不会发生变化。

6. 在计算这些概率的时候，可以先求出事件在一年内未发生的概率。例如，如果事件的概率为1/2，那么这个事件在一年内不发生的概率等于（0.999）^{365}＝0.69。因此，该事件发生的概率等于31%。一个概率为一百万分之一的事件在一个世纪内不发生的概率也可以用类似的方法计算。

7. Cederman, 2003; Clauset、Young&Gleditsch, 2007; Roberts&Turcotte, 1998.恐怖袭击导致x人死亡的概率大体上可以写为一个常数项：0.06除以x的平方。对于x只取整数值的离散分布，幂律分布则可以写为$p(x) = 0.608x^{-2}$。之所以选择系数0.608，是为了使概率总和等于1。因为$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1.644934$；0.608与1.644934的乘积等于1。

8. 对于幂律，我们对两边同时取对数，并将$y = Cx^{-a}$变换为$\log(y) = \log(C) - a\log(x)$，这是一个$\log(y)$关于$\log(x)$的线性方程。当我们绘制出$\log(y)$随$\log(x)$而变的图形时，将会得到一条直线。对于指数分布，$y = C \times A^{-x}$。如果取两边的对数，就可以得到$\log(y) = \log(C) - x\log(A)$，这意味着$\log(y)$对于$x$是线性的。换句话说，$\log(y)$将随$\log(x)$而快速减少，从而形成凹的图形。

9. 如果我们取对数正态分布的对数，就会得到以下形式的方程：$\log(y) = C - b \times \log(x) - \frac{\log(x)^2}{\sigma^2}$，其中，$\sigma$是这个对数正态分布的标准偏差的自然对数，作为该分布方差的一个代理变量。对于很大的σ，$\log(x)^2$的贡献将很小，直到$\log(x)$的值变得足够大，可以引起图形出现下降为止。

10. 关于对数正态分布和幂律分布的正式区别，请参见：Broido&Clauset, 2018。研究认为，许多曾经被认为具有幂律分布特点的网络，实际上并不是幂律分布的。

11. 要了解齐普夫的单词频率定律和其他可选模型，请参见：Piantadosi, 2014。如果事件的大小分布满足幂律，那么排名也符合这个定律。在开放区间［1，∞）上，指数为a的幂律分布具有形式$p_a(x) = ax^{-a}$。假设有100个事件。令S_R表示第R大的事件的预期大小。那么某个事件大于S_R的概率必定等于$\frac{R}{100}$。例如，如果$R = 3$，那么一个事件大于S_3的概率必定等于3%。因此：

$$\int_{S_R}^{\infty} ax^{-a} = \frac{R}{100}$$

从中可以解得$S_R^{-(a-1)} = \frac{R}{100}$，而它又可以改写为：

$$S_R = \left(\frac{R}{100}\right)^{\frac{-1}{(a-1)}}$$

。

在$a=2$的特殊情况下，上述表达式就变成了$S_R = \left(\frac{R}{100}\right)^{-1}$。

12. Bak, 1996.这个模型的适用范围很广，至今仍然没有看到它的边界。学者们已经利用这个模型来解释经济波动、战争中的死伤、恐怖主义行为、进化中的间断均衡以及交通拥塞等理论和现实问题了。请参见：Paczuski&Nagel, 1996；Sneppen et al., 1995。

13. Salganik、Dodd&Watts, 2006; Ormerod, 2012.幂律分布还意味着，构成概率分布的很大一部分是非常多的小事件。这些小事件合到一起，也足以产生与大事件同等规模的经济价值。请参见：Anderson, 2008b。互联网的出现，使零售商能够出售大量的书籍、电影和音乐，尽管有些东西只能吸引少数人。一个只卖一本畅销书、卖出了500万的零售商，获得的利润可能与另一个销售1万种书、每种书只卖出500本的零售商差不多。

14. 有一个模型说明了这种情况是怎样出现的，请参见：Denrell&Liu, 2012。

15. 地质学家使用里氏标准来表示地震震级。震级等于地震强度的对数。因此，里氏6级地震的强度是里氏5级地震的10倍。地震的强烈可以用齐普夫定律来预测（但不是地震的时间），请参见：Merriam&Davis, 2009。

16. Eliot、Golub&Jackson, 2014.他们给出了增大连通性如何使失败的可能性降低的明确模型。

17. Mayt、Levin&Sugihara, 2008.

18. Stock&Watson, 2003.

19. Carvalho&Gabaix, 2003.

20. Claridat、Galí&Gertler, 2000.

21. 我要感谢塞思·劳埃德（Seth Lloyd），是他提醒我注意到了这个例子。

22. 我们让工资分布等于随机变量x的100 000倍，其中，$p(x)=2x^{-3}$从1到∞。变量x的均值等于2，因此工资分布的均值为200 000美元。

23. Weitzman, 1979.他给出的模型在更一般的情况下证明了这个结果。

24. Bell et al., 2018.

07 线性模型

1. 有价值的葡萄酒，如波尔多葡萄酒，都要由专家加以排名。这些葡萄酒同时也在市场上定价。价格和排名都可以作为质量的代理变量。奥利·阿申费尔特（Orley Ashenfelter）根据冬季降雨量、收获季降雨量和9月份的平均温度，拟合出了关于波尔多葡萄酒的质量的（对数）线性模型。请参见：Ashenfelter, 2010。线性模型将因变量的对数表示为自变量的对数的线性和，例如：

$$\log(y)=b_0+b_1\log(x_1)+b_2\log(x_2)$$

这个表达式意味着，因变量也可以写成自变量的乘积。只需对上式的两边同时取以e为底的指数，就可以做到这一点。于是，我们得到：

$$y=e^{b_0}(x_1)^{b_1}(x_2)^{b_2}$$

通过取对数，乘式就可以变加式，然后就可以应用线性回归的工具了。以年份葡萄酒的价格为因变量，阿申费尔特的模型的拟合度指标（R^2，即，得到了解释变差的百分比）达到了83%左右。有证据表明，它比葡萄酒专家的判断更加准确地预测了葡萄酒的价格。阿申费尔特的模型甚至可以预测专家的估值的变化。罗伯特·帕克（Robert Parker），一位著名的葡萄酒评酒专家，最初将Pomerol和St. Emilion于1975年出品的葡萄酒评为95分（满分为100分）。但是，在阿申费尔特的模型的预测中，对这些酒的质量排名则较低。后来，到了1983年，罗伯特·帕克将他对这些酒的排名降低到了低于平均水平，一如阿申费尔特的模型的预测。请参见：Storchmann, 2011。

2. Xie, 2007.

3. Ryall&Bramson, 2013.

4. Mauboussin，2012.他讨论了如何利用这个方程式指导合理的管理决策。

5. Bertrand&Mullainathan, 2001.

6. Shapiro、Meschede&Osoro, 2013.他们并没有把相关性和因果关系混为一谈。如果两个变量不相关，就不应该期待它们之间存在因果关系。

7. 为了找到最优的那条线，我们使用的是支持向量机（SVM），这是一种类似于回归的方法。两者之间关键区别在于，支持向量机将所有数据分成正的组或负的组，然后找到最大化到每组中最近点的直线。如果不存在这样的直线，就会对偏离行为进行处罚。回归则考虑到所有点的距离。

08　非线性模型

1. Arthur, 1994.

2. 30次倍增为2^{30}，超过10亿。

3. Karlsson, 2016.

4. Ebbinghaus, 1885.

5. 关注大脑反应的研究者发现，即便是享用巧克力，消费了一定数量后，人们也会开始觉得厌恶，请参见：Small et al., 2001。

6. 与本书中的许多例子一样，这个例子我也是从拉夫和马奇出版于1975年的那本书中借用来的。

7. 假设你每年投资3 000美元。如果股票的价格维持在每股15美元的水平不变，那么你每年可以买200股。如果价格在20美元到10美元之间变动，那么在价格最高的年份你只能购买150股，而在价格较低的年份你可以购买300股。平均而言，你可以购买225股，这比你在价格不变的情况下购买的数量更多。

8. 关键假设是，在柯布-道格拉斯模型中的指数a和（$1-a$）的总和为1。这个假设意味着如果我们将工人数量和资本数量加倍，总产出也会

增加一倍，即：

产出＝常数×（2×工人）^{a}（2×资本）$^{(1-a)}$

展开各项并重新排列，就可以得到输出加倍的方程式：

产出＝常数×（2×工人）^{a}（2×资本）$^{(1-a)}$

9. 在第二年，每天产出等于$100\sqrt{2}=141$。在第三年，产出则$100\sqrt{3}=173$。增长率等于每年的产出增长百分比。

10. 计算过程如下。第2年，机器：290；产出：1 702（$=100\sqrt{290}$）。投资＝（0.2）×1 702＝340，因此消费＝1 362。折旧＝（0.1）×290＝29。第3年，机器：601，产出：2 453（$=100\sqrt{601}$）。投资＝（0.2）×2 453＝491，因此消费＝1 962。折旧＝（0.1）×601＝60。

11. 长期均衡可以通过求解M^*求出，即，使得$0.2\times 100\sqrt{M^*}=0.1\ M^*$。$M^*$＝40 000。

12. 在完整的索洛增长模型中，会用一个参数a替换掉这里的平方根函数，像在柯布-道格拉斯模型中一样，还包括劳动市场。

13. 为了求出均衡，我们设定投资等于折旧：$s\times A\sqrt{L}\sqrt{K^*}=d\times K^*$。因此，均衡机器数量$K^*$要满足等式$\sqrt{K^*}=A\frac{s}{d}\sqrt{L}$。将之代回产出函数，可以求得产出等于$A^2L\frac{s}{d}$。

14. 经济学家罗伯特·戈登（Robert Gordon, 2016）则采取了一种悲观的看法，认为即将出现的新技术将使得A出现大幅增长。更高级的增长模型将技术视为其他变量的函数。例如，保罗·罗默（Paul Romer, 1986）的模型假设增长源于商品种类的增加：随着经济的增长，产品种类也在增长。经济学家马丁·魏茨曼（Martin Weitzman, 1998）则对思想的产生和组合给出了明确的模型。

15. 关于技术变革中的滞后，请参见：Arthur, 2011。

16. 例如，以前从来没有通过固定电话的乡村地区，可以直接建造无线基部并提供手机服务。亚历山大·格申克伦（Alexander Gerschenkron, 1952）将这种现象称为“后发优势”。

17. 皮凯蒂（2014）揭示，从1700年到2012年，世界GDP的年平均增长率仅为1.6%，而且这种增长中的一半是来自人口增长。将72法则应用于这0.8%的年增长率，我们可以发现，在300年的时间内，平均

生活水平提高了大约10倍。

09　与价值和权力有关的模型

1. 这个值是这样计算出来的：将桨手能够增加价值的概率1/6乘以桨手能够增加的价值的期望值（后者等于4/5×10再加1/5×2）。同时注意到，夏普利值的总和为10，也就是这个博弈的总价值。

2. 正式的计算过程如下：在阿伦的6个想法中，只有两个是他独有的，一个是贝蒂也提出来的，3个是每个人都提出来的。如果他是最先进入的那个人（这个事件发生的概率为1/3），那么增加的价值为6。如果他是第二个进入的，那么能够增加两个独有的想法，这两个想法都是其他人所不知道的；而且，在贝蒂和卡洛斯进入之前，他还有1/2的机会再增加一个想法。所以，他总共增加了2.5个想法。如果他是第三个进入的，那么他可以增加两个他独有的想法。因此，他的夏普利值等于3.5。贝蒂提出了4个与另外一个人共同的想法、3个与所有人共同的想法。因此，她的夏普利值等于3。最后，卡洛斯提出了3个与另一个人共同的想法以及3个与所有人共同的想法，所以他的夏普利值等于2.5。需要特别注意的是，夏普利值的总和为9，即所有想法的总数。

3. 衡量投票权力的另一种方法是，班茨哈夫-彭罗斯指数（Banzhaf-Penrose index）。它的计算方法是，根据所有可能的获胜联盟确定所有可能成为“关键投票者”的政党的数量，而赋予一个政党的值就等于它发挥关键作用的次数的值除以这个数字。请参见：Banzhaf, 1965。

4. Groseclose&Snyder, 1996.

10　网络模型

1. 对网络模型的全面讨论，请参见：Newman, 2010；想要了解网络模型在社会科学领域的应用，请参见：Jackson, 2008；Tassier, 2013。

2. 给定任何一个节点，它的最小路径长度可以为1到4个节点，也可以为2到4个节点，还可以为3到4个节点，由此，从每个节点开始的最小

路径共有12个节点。平均而言，节点的最小路径是访问另一个节点。由此可以得出，每个节点的平均介数等于1/12。根据对称性，所有节点必定都具有相同的介数。

3. 关于社区探寻算法的更多细节，请参见：Newman，2010。由于可能性的数量不同，用这些算法构造出来的分区可能有所不同。具有100个节点的网络可以用超过1.9亿个不同的方法进行分区。删除边的次序的随机性，相同的算法通常也会产生不同的分区。通过应用多个算法并多次应用每个算法，就可以增强推理的鲁棒性。

4. 根据中心极限定理，我们知道度是正态分布的，并且其均值为 $\frac{2E}{N}$ ，因为每个边连接两个节点。

5. Watts&Strogatz, 1998.

6. 关于网络形成的正式分析，请参见：Newman, 2010。

7. Ugander et al., 2011。

8. 给定一个具有N个人的网络，令d_i表示节点i的邻居的数量，即度。平均度 $\bar{d}$ 可以表示为如下形式：

$$\bar{d}=\frac{\sum_{i=1}^{N}d_i}{N}$$

平均度 $\bar{d}$ 等于这个节点的期望邻居数。而在计算邻居的邻居的平均数量时，度为d_i的节点会被“数上”d_i次，即，每个邻居一次。因此，节点的邻居的邻居的总数N_2可以表示如下：

$$N_2=\sum_{i=1}^{N}d_i^2$$

因此，为了求得一个节点的邻居的平均度，我们必须用这个数除以邻居的总数，得出的结果为$N\bar{d}$。因此，只要证明下式就足够了：

$$\frac{\sum_{i=1}^{N}d_i^2}{N\bar{d}}\geqslant\bar{d}$$

而它可以重写为

$$\frac{\sum_{i=1}^{N}d_i^2}{N}-\bar{d}^2\geqslant 0$$

左边的项等于度分布的方差。如果任意两个节点具有不同的度数，那么度分布就具有正的方差，因此，节点的邻居的平均度，超过节点的平均度。

9. Eom&Jo, 2014.

10. Dodds、Muhamad&Watts, 2003.

11. Newman, 2010; Jackson, 2008.

12. Granovetter, 1973.

13. 四度朋友的数量是通过对如下8组节点求和来计算的：CRCR＝4 000 000，CRRC＝4 000 000，RCRC＝4 000 000，CRRR＝800 000，RCRR＝800 000，RRCR＝800 000，RRRC＝800 000，RRRR＝160 000。

14. Albert、Albert&Nakarado, 2004.

15. Groysberg, 2012．条纹模型可以解释它们为何无法成功回归均值。

16. 关于填补结构洞的人的价值，请参见：Burt, 1995。

17. 关于对教师网络的影响，请参见：Frank et al., 2018。

18. 具有非零价值的联盟是{A，B}、{B，C}和{A，B，C}。在前两个联盟中，每个联盟的价值都等于10。第三个联盟的附加价值等于-6。第三个联盟的独立价值等于14，但是这个联盟是由两个联盟组成的，每个联盟的价值都是10。因此，联盟的价值等于14减去20：-6＝（14-10-10）。然后，我们可以为每个联盟中的每个博弈参与者分配如下夏普利值：联盟{1，2}：博弈参与者1为5，博弈参与者2为5；联盟{2，3}：博弈参与者2为5，博弈参与者3为5；联盟{1，2，3}：博弈参与者1为-2，博弈参与者2为-2，博弈参与者3为-2。对这些值求和，就可以得到迈尔森值。

11　广播模型、扩散模型和传染模型

1. 这些模型都假设了离散的时间（步长），例如，天或周，并使用差分方程来描述“明天”被感染（或变得知情）的人的数量与“今天”被感染（或变得知情）的人的数量之间关系。连续时间模型则需要运用微分方程和微积分。但是，如果改用连续时间模型，那么结果从定性的角度来看不会有所不同。

2. 将第一个方程式插入第二个方程式，可以得出：36 000＝20 000＋20 000-P_{broad}×20 000，进一步化简为4 000＝P_{broad}×20 000，所以我们有：P_{broad}＝0.2，以及N_{POP}＝100 000。

3. Griliches, 1988.

4. 假设，每个应用程序的初始销售量（I）等于100。一开始，设定$P_{diffuse}$＝0.4，同时POP＝1 000。第三期的新销售量等于$0.4 \times \frac{100 \times 900}{1\ 000} = 36$。与此类似，我们可以计算出未来各期的销售数据。对于第二组数据，令$P_{diffuse}$＝0.3，POP＝100 000。第二期的新销售量$0.3 \times \frac{100 \times 999\ 900}{1\ 000\ 000} = 30$。用同样的方法，可以求得后续各期的销售量。

5. 巴斯把率先采用某种技术或购买某种产品的人称为“创新者”，而把跟着这样做的人称为“模仿者”（Bass, 1969）。

6. 对R_0的正式推导始于这样一个观察结果：当感染者的人数很小的时候，易感者的数量大体上相当于相关人群的数量。因此，为了减少变量的数量，我们可以用相关人群的数量代替易感者的数量。然后，我们可以将感染者人数的变化写成最初感染者人数的线性函数的形式（见正文中的专栏）。对R_0的正式推导过程如下。当出现了一种新的传染病时，它最初只会感染少数人。我们用I_0表示这个很小的感染者的数量。将之代入SIR模型，第一期中感染者的数量等于：

$$I_1 = I_0 + P_{contact} \times P_{spread} \times \frac{I_0}{N_{POP}} \times S_0 - P_{recover} I_t$$

以N_{POP}为S_0的近似，上面这个式子就变为$I1 = I0 + P_{contact} \times P_{spread} \times I1 - P_{recover} I1$。由此我们可以推出，感染者的人数将上升，当且仅当$P_{contact} \times P_{spread} > P_{recover}$，而这又等价于下式：

$$\frac{P_{contact} \times P_{spread}}{P_{recover}} \geqslant 1$$

7. 强制隔离会使接触的概率降低到接近零的水平，从而有效降低基本再生数，但是成本很高。在20世纪初，结核病（$R_0 \approx 3$）在美国每年都会导致超过10万人死亡，各州都不得不加征财产税，以建立强制

性的疗养院以收容隔离病人，因为手术治疗如摘除肺叶和整个肺，甚至往肺部空洞处塞入乒乓球，但最终全都被证明是无效的。请参见：Dubos, 1987。

8. 为了求疫苗接种阈值，我们需要对SIR模型进行一些修正，以便容纳接种了疫苗的人群。一种传染病要得到传播，感染者必须与未接种疫苗的人（概率为（1-V））接触（接触概率为$P_{contact}$），并且传染病必须是会扩散的（P_{spread}）。第一期的差分方程可以写为如下形式：

$$I_1 = I_0 + P_{contact} \times P_{spread} \times \frac{I_0}{N_{POP}} \times (1 - V) \times S_0 - P_{recover} I_t$$

运用近似值$S_0 = N_{POP}$（就像在推导R_0时一样），上式就可以改写为：

$$I_1 = I_0 + P_{contact} \times P_{spread} \times I_0 \times (1-V) - P_{recover} I_t$$

当且仅当$P_{contact} \times P_{spread} \times (1-V) > P_{recover} I_t$时，感染者的人数才会增加。这个不等式可以重写为$R_0$（1-V）≤1。展开并重新排列各项，我们有$R_0 - 1 < V \times R_0$。只要两边同时除以$R_0$就可以得到我们想要的结果。

9. 关于从社会科学角度对SIR模型和群体免疫力的更详尽的分析，请参见：Tassier, 2013。

10. Stein, 2011.

11. Updike, 1960.

12. Tweedle&Smith, 2012.

13. Lamberson&Page, 2012b.

14. 数据源自网络。

15. Christakis&Fowler, 2009.

16. 对此，达蒙·森托拉（Damon Centola）和迈克尔·梅西（Michael Macy）（2007）称之为复杂的传染病（complex contagion）。

12 熵：对不确定性建模

1. Smaldino, 2013.

2. 数字x的以2为底的对数必定等于使2产生x的幂，因此$\log_2(4)=2$，且$\log_2(2^N)=N$。在一般情况下，$\log_a(x)$等于使a产生x的幂。因此，如果$a^y=x$，则$\log_a(x)=\log_a(a^y)=y$。

3. 我们可以将信息熵写成如下的长形式：

$$H=-\frac{1}{2}\times\log_2(\frac{1}{2})-\frac{1}{8}\times\log_2(\frac{1}{8})-\frac{1}{8}\times\log_2(\frac{1}{8})-\frac{1}{8}\times\log_2(\frac{1}{8})-\frac{1}{8}\times\log_2(\frac{1}{8})$$

这个式子可以化简为：

$$\frac{1}{2}(1)+\frac{1}{8}(3)+\frac{1}{8}(3)+\frac{1}{8}(3)+\frac{1}{8}(3)=\frac{1}{2}(1+3)=2$$

4. 多样性指数，即，概率的平方和的倒数，可以满足前两个公理以及乘法公理。以多样性指数来衡量，已知结果的多样性等于1。请参见：Page, 2007、2010a。

5. Wolfram, 2001; Page, 2010a.

6. 亚历山大总共列出了15个这样的属性。他将自己的想法，再加上很多精英的照片集结成册，自行出版了4本精美的书：《秩序的本质，第1册：生命现象》（*The Nature of Order, Book 1: The Phenomenon of Life*）（2002）、《秩序的本质，第2册：创造生命的过程》（*The Nature of Order, Book 2: The Process of Creating Life*）（2002）、《秩序的本质，第3册：生活世界的愿景》（*The Nature of Order, Book 3: A Vision of a Living World*）（2005），以及《秩序的本质，第4册：流光溢彩的大地》（*The Nature of Order, Book 4: The Luminous Ground*）（2004）。这4本书中，第2册与我们这里讨论的主题的关系最密切。

13　随机游走

1. 欲了解更多关于随机游走的知识，请参见：Mlodinow, 2009。

2. Taleb, 2001.

3. Turchin, 1998; Suki&Frey, 2017.

4. 需要注意的是，大数定律说明平均比例是收敛的，而中心极限定理则

告诉我们白球比例的分布是正态的。

5. 一位三分球命中率为46%的篮球运动员连续投中9个三分球的概率大约为1/1 000（0.46^9）。如果这个篮球运动员一直继续投三分，那么在10年的NBA生涯中（总共参加大约800场比赛），一次都做不到三分球9投9中的概率大约为47%（0.999^{800}）。

30多年来，统计学家一直在争辩，篮球运动员和其他职业运动员是不是真的会出现“热手效应”，或者换句话说，任何一次投篮或罚球命中的概率是否与前一次的成功相关。请参见：Chance, 2009。他分析了乔·迪马吉奥（Joe DiMaggio）的56场比赛连胜纪录。在考虑“热手效应”的证据时，我们必须考虑到行为。如果一名篮球运动员认为自己的手感很好，那么可能会尝试更加困难的投篮。此外，如果防守方觉得对方哪个运动员手感太好，他们就会收紧对他的防守。这些行为反应在模型中可以通过加大难度来体现。特沃斯基等人没有发现可以证明“热手效应”的证据。乔舒亚·米勒（Joshua Miller）和亚当·圣胡尔霍（Adam Sanjurjo）（2015）则在以往的研究中的条件概率计算中发现了一个推理错误，进而证明以前那些声称没有发现“热手效应”的研究实际上支持了“热手效应”假说。以前研究的这种误差源于采样技术中的缺陷。这些研究收集了多名（例如）篮球运动员投篮命中和不中的序列；然后再计算随机选择的多次投篮后某次投篮命中的概率。这种采样程序会引入一种微妙的统计偏差。

为了说明这种偏差，不妨考虑如下这种情况：多名篮球运动员每人只投篮4次，而且每次投篮都有相同的命中概率，这样就有16种可能的投篮命中和未命中的序列。我们用B代表投篮命中，用M代表投篮不中。在这16个序列中，有6个是连续两次命中后再接着投篮一次产生的，它们是：BBBB，BBBM，MBBB，BBMB，BBMM和MBBM。这些都是连续两次投篮命中后再投第三次篮的情形。如果抽到了BBBB，那么无论选择哪一个由两个B组成的序列，那么接下来再投篮命中的概率都等于100%。如果选择了MBBB，那么B跟着BB出现的概率也等于100%。但是，如果选择了BBBM，那么BB之后出现M的概率等于50%。最后，如果选择的是MBBM、BBMB或BBMM，那么M必定会在BB之后出现。求这6种情况的平均值：

$$p(\mathrm{B}|\mathrm{BB})=\frac{1}{6}\left[1+1+\frac{1}{2}+0+0+0\right]=\frac{5}{12}$$

之所以会产生偏差，是因为在序列BBBB中，有两个BB可以选择，但是在其他序列中，则只有一个（例如，BBMB）。前述采样程序使得BBBB的这两个子序列中的每一个被选中的概率都仅为BBMB中的那一个子序列的一半。这种偏差的含义是，如果不存在“热手效应”，那么这种采样程序就已经证明投篮命中后再投篮时更有可能不中。当然事实并非如此，而这就意味着投篮命中之后再投篮实际上更有可能命中。

6. 麦道夫几十年来一直宣称自己的客户每月的回报率为1.5%。他声称，无论市场大环境（“大盘”）如何变化，他的投资每个月都在升值。在经济衰退期间维持正收益，要比表现优于市场更加困难。当市场下跌时，一个人即使战胜了市场也仍然可能会出现负收益。整体市场下跌了80多个月，但是麦道夫却一直声称自己实现了正收益。如果我们愿意冒风险做出一个“英勇”的假设：即使在大盘下跌的时间超过3/4的情况下，麦道夫也能以某种神奇的方法取得正收益，那么他连续80期成功的概率大约为十亿分之一（0.75^{80}）。

7. 只要了解到，一个随机游走的值是相同的、独立的随机变量的总和，我们就可以求解标准差。每个随机变量的均值都为零，而且取值为+1或-1。因此，每个都具有1的标准差。设置$\sigma=1$并且对所得到的总和应用西格玛平方根公式，就可以得出结果。

8. Newman, 2005.

9. Levinthal, 1991; Axtell, 2001.

10. Newman, 2005; Sneppen et al., 1995.

11. 湖泊是按面积计算的。我们的模型产生幂律直径。面积等于直径平方的常数倍，因此面积也将是幂律分布的。具体数据请参见：Downing et al., 2006。

12. 在平衡的轮盘赌轮上，所有间隔都应有相同的可能性。如果桌子有任何倾斜，那么球在上坡时更可能从外缘掉落。关于多恩·法默、诺曼·帕卡德（Norman Packard）和他们的朋友通过“可穿戴计算机”，利用这个现象并击败轮盘赌的故事，请参见网络。

13. N次投注后，这随机游走的期望值等于$N\times\left(\frac{18}{38}-\frac{20}{38}\right)=\frac{-N}{19}$。由于获胜的概率大约为$\frac{1}{2}$，我们可以将这个值的标准偏差近似为$\sqrt{N}$

。确切的值则等于$2\sqrt{\frac{18}{38}\times\frac{20}{38}N}=0.9986\sqrt{N}$。

14. L.皮尔（L. Peel）和A.克劳塞特（A. Clauset）（2015）将每个博弈模型化为一个单个序列，并发现得分序列表现出了反持续性：上次得过分的队伍不太可能接下来马上得分。鉴于球队交替发动进攻，这个结果应该是预料之中的。

15. Baxter, 2009.

16. 这个结果的证明需要先计算在N个步骤内返回起点的概率，然后把所有可能的N的结果都加总起来。

17. Samuelson, 1965.

18. Grossman&Stiglitz, 1980.

19. Lo&MacKinlay, 2007.席勒（2005）证明，市盈率相对较低的股票的表现优于市场。

20. Mauboussin, 2012.

21. 从1967年到2017年的这个窗口期恰好正逢标准普尔500指数上涨周期。在大多数窗口期，股票价格增长速度高于经济整体的增长速度。

14　路径依赖模型

1. Hathaway, 2001.

2. Pierson, 2004.

3. Bednar&Page, 2007、2018.

4. Page, 2006.

5. 这里是我在2006年给出的一个证明概要。显然，只要证明在前N期抽取出K个白球的概率等于1/（N+1）即可。存在（N+1）种可能性，因为K最小可以取零值，最大则可以等于N。N个周期中抽取出给定K个白球序列的概率可以写为N个分数的乘积。这些分数的分母分别是数字2到N+1。分子分别为数字1到K（白球）和1到（N-K）（灰球）。分子的乘积等于K！乘以（N-K）！而分母的乘积则等于（N+1）！。通过这个计算，就可以给出K个白球的特定序列的概率。在N个期内排列K个球的可能方式的数量等于

$\binom{N}{K}=\frac{N!}{K!(N-K)!}$。因此，正好抽取出$K$个白球的总概率等于

$$\frac{K!(N-K)!}{(N+1)!}\cdot\frac{N!}{K!(N-K)!}=\frac{1}{(N+1)}$$

6. 我们用反证法来证明，假设结果不成立。从长期看，60%的结果是白色的，因此，瓮中60%的球将是灰色的。但是，这又意味着60%的结果将是灰色的。于是，出现矛盾，原命题得证。

7. 要想了解更多依赖用熵来度量不确定性的模型，请参见：Lamberson&Page, 2012b。

8. Lamberson&Pag, 2012b.

9. Page, 1997.

10. 风险价值也可以用年内任何时候损失超过10 000美元的概率来计算。

11. 这些计算依据的是以下事实：长度为N的随机游走的值的标准差等于$\sqrt{N}$。2.5%对应于两个标准差。

15　局部互动模型

1. 在物理学中，局部多数模型通常称为伊辛模型（Ising model）。局部多数模型是投票者模型的一个变体。投票者模型假设了随机选择各种规模的邻居。请参见：Castellano、Fortunato&Loreto, 2009。

2. 对于沿4条边的元胞，我们将顶部的边连接到底部的边，创建出一个圆柱体，然后将左侧的边连接到右侧的边，创建出一个圆环。

3. 在局部多数模型的另一个版本中，元胞可以被同时激活，或者根据更新的动机来激活，其中处于相反状态的最大多数邻居元胞最先移动。如果元胞同时更新，局部多数模型可以生成循环。

4. 关于起立鼓掌的模型，请参见：Miller&Page, 2004。

5. 关于包括了跨域行动一致性的文化模型，请参见：Bednar et al., 2010。

6. 在这里，也是先将顶部连接到底部创建出一个圆柱体，然后连接两端以形成一个“瘦小”的圆环。

7. 有一类称为扩散反应的模型也会在较窄的形状上产生条纹，并在较宽的形状上产生斑点。使用这些模型，科学家可以预测哪些动物会有条纹，哪些动物会有斑点，哪些动物会是纯色。答案取决于在发育阶段的哺乳动物胚胎的大小，即在模式形成时的大小，而不是成年动物的大小。否则，大象就会有斑点。请参见：Murray, 1988。

8. 我要感谢伯纳多·休伯曼（Bernardo Huberman）提供给我这个比喻。

9. 要证明这一点，需要编写一个计算机程序。这个程序能够生成随机数字序列或随机模式，并证明生命游戏模仿了计算机程序。正式的证明需要说明生命游戏在元胞自动机的集合中是通用的。请参见：Berlekamp、Conway&Guy, 1982。

10. Dennett, 1991; Hawking&Mlodinow, 2011.

16　李雅普诺夫函数与均衡

1. Nagel, 1995.

2. 我要感谢珍娜·贝德纳，她提供了关于逐底竞争的几个例子，除此之外，她还提供了本书的很多个例子。

3. Page, 2001.

4. 外部性（即便是负外部性）的存在，不能排除构建李雅普诺夫函数的需要。局部多数模型和路径选择模型都包含了负外部性。在局部多数模型中，当一个元胞改变了自己状态时，它会对处于相反状态的邻居施加负外部性。然而，它与现在匹配的邻居产生的正外部性更大。

5. Guy, 1983.如果你够大胆，那么就从27开始试一试。

17　马尔可夫模型

1. 正式的研究论文需要利用统计技术更准确地估计转移概率并计算出误差范围。它还要检验转移概率在那个时间段内是不是保持不变的。如果转移概率取决于人均收入，那么就不会。请参见：Przeworski et al., 2000。

2. Flores&Nooruddin, 2016.

3. Tilly, 1998.

4. 采用瓷砖的人的均衡百分比等于从油毡转移到瓷砖的概率除以从油毡转移到瓷砖的概率与从瓷砖转移到油毡的概率的总和。这就等于

$$\frac{\frac{1}{4}\cdot\frac{1}{10}}{\frac{1}{4}\cdot\frac{1}{10}+\frac{3}{4}\cdot\frac{1}{60}}=\frac{\frac{1}{40}}{\frac{3}{80}}=\frac{2}{3}$$

。我们还可以写出如下更一般的模型。令D表示拥有昂贵的耐用品的人的百分比，C表示拥有更便宜的商品的人的百分比。再令$\mathrm{BUY}(\mathrm{C})>\frac{1}{2}$表示有人购买更便宜的商品的概率。令REPLACE（C）和REPLACE（D）表示更换两种产品的概率。如果以下不等式成立：

REPLACE（C）×1-BUY（C）＞REPLACE（D）×BUY（C）

那么会有更多的人买更便宜的商品，但是同时更多的人拥有耐用品。根据假设，命题的第一部分显然成立。为了证明第二部分，我们要先解出转移概率。某个人从D移动到C的概率是*P*（D，C）＝REPLACE（D）×BUY（C）。与此类似，某个人从C移动到D的概率是*P*（C，D）＝REPLACE（C）×1-BUY（C）。在均衡状态下，D-D×P（D，C）＋C*P*（C，D）＝D。设定C＝（1-D），我们得到

$$\mathrm{D}=\frac{P(\mathrm{C},\mathrm{D})}{P(\mathrm{C},\mathrm{D})+P(\mathrm{D},\mathrm{C})}$$

。如果*P*（C，D）＞*P*（D，C），那么上式大于0.5。这个不等式与REPLACE（C）×1-BUY（C）＞REPLACE（D）×BUY（C）是等价的。

5. McPhee, 1963; Ehrenberg, 1969.他们给出了一些证明双重危险的经验证据。为了证明这个结果，我们只需要证明消费者在转换品牌时会以同样的概率购买任何一种产品。

6. Briggs&Sculpher, 1998.

7. Schrodt, 1998.

8. Khmelev&Tweedie, 2001.

9. Khmelev&Tweedie, 2001.

10. Reynolds&Saxonhouse, 1995.

11. 请参见网络。

12. 巧妙构建此类模型的关键在于，定义有用的状态并分配准确的转移概率。请参见：Langville&Meyer, 2012。

13. 在食物链中，一个物种与它所消费的物种相连。请参见：Allesina&Pascual, 2009。

14. Russakoff, 2015.

18 系统动力学模型

1. Sterman, 2000.

2. 关于定性系统动力学模型的价值分析，请参见：Wellman, 1990。

3. 这个模型假设所有野兔都是因被狐狸吃掉而死亡的。为野兔的其他死亡方式增加变量，只会使模型复杂化，但不会改变其结果，因为这样做其实只降低了现在这个模型中野兔增长率的取值范围。这里的符号$\dot{H}$表示每单位时间内H的变化率，换一种写法是$\frac{\partial H}{\partial t}$。当野兔和狐狸数量的变化率都等于零时，即$\dot{H}=\dot{F}=0$时，均衡就实现了。为了求解均衡，取方程$gH-aFH=0$，然后两边同时除以$H$，可以得到$g-aF=0$。解出$F$就可以得到结果。接下来，再取方程$bFH-dF=0$，然后两边同时除以$F$，可以得到$bH-d-0$。求解$H$就可以得到结果。

4. 我要感谢多伦多大学的迈克尔·赖亚尔（Michael Ryall），他提供了这个例子。

5. 对于这个模型的相关结果的总结，请参见：Meadows et al., 1972。

6. 这套模型就是通常人们所称的“罗马俱乐部”模型。戴维·洛克菲勒（David Rockefeller）于1968年创立的罗马俱乐部，资助了围绕这些模型的研究，并推广了这些模型的发现。

7. 通过对变量在更小范围内进行操纵，约翰·米勒还可以将模型的预测推向近300亿。请参见：Miller, 1998。

8. Hecht, 2008；MacKenzie, 2012.

9. Sterman, 2006.

10. Glantz, 2008.

19 基于阈值的模型

1. Granovetter, 1978.

2. 爱彼迎创始人采取的方法是，在2008年美国总统大选期间，开发出了奥巴马奥氏（ObamaO's）和船长麦凯恩氏（Cap'n McCain's）谷物早餐，然后挨家挨户上门推销。

3. 关于旋转门模型，请参见：Jacobs, 1989。实证研究表明，在几乎不需要接受传统"正规"教育的职业中（例如调酒行业和园艺行业），当某个工作场中女性员工的比例占到了15%时，男性就会选择离开（或选择不进入），请参见：Pan, 2015。

4. Syverson, 2007.

5. Gammill&Marsh, 1988.

6. Easley et al., 2012.

20 空间竞争模型与享受竞争模型

1. Clark、Golder&Golder, 2008.

2. 关于对法律立场的分析，请参见：Martin&Quinn, 2002。

3. 霍特林（1929）研究了地理区位决策。凯尔文·兰开斯特（Kelvin Lancaster）（1966）扩展了霍特林的模型，用来研究经济选择之间的享受竞争；唐斯（1957）则将这个模型应用于政治学研究。

4. 提名分数模型提供了一种更复杂的方法来分配意识形态立场，但是所依据的基本思想是一样的，请参见：Poole&Rosenthal, 1985。

5. 选择常数项（表达式中的C）时，要让所有收益均为正。为了实现这一点，可以将它设定为任何理想点和可选点之间的最大可能距离。

6. 在以这种方式构建分割线时，我们假设消费者赋予这两个属性的权重是相同的，也可以将权重不同的情况包括进来。如果人们认为含糖量比可可含量更重要，那么我们会以逆时针的角度倾斜这条分割线。而在极端情况下，如果（例如）人们只重视含糖量一个属性，那么分割线将是水平的，并且会在垂直轴上均匀地划分A和B。空间模型提供了一个清晰的例子，说明我们应该怎样将直觉——在这里直觉是我们更喜欢且更接近我们的理想的东西，转换为正式的模型。一旦将备选方案（巧克力棒）和消费者的理想点在空间上画出来，并根据它们与理想点之间的距离来定义偏好排序（将备选方案

从最优到最差排列好），我们事实上就已经根据这些备选方案给出了一个效用函数了。一种产品的效用等于它与理想点之间的距离的倒数。

7. Havel, 1978.

8. Martin&Quinn, 2002.

9. McCarty，2011.这些百分比可能会随着对党派忠诚度的变化或者提交投票的法案类型的变化而改变。

10. 从形式上说，如果我们假设选民的人数为奇数并且在二维中位数上存在一个选民理想点，那么这个条件就要求通过二维中位数的任何一条线都会将剩下的选民理想点划分为两个大小相等的集合。请参见：Plott, 1967。

11. 理查德·麦凯尔维（Richard McKelvey，1979）证明，连续进行一系列投票，会导致具有两个维度或多个维度的任何一个政策都有可能会被选中，从而导致“混乱结果”的发生。不过，麦凯尔维谨慎地声明，他这个结论只针对给定偏好下可能产生的结果序列，而不是对系列投票会发生什么事情的预测。也请参见：Kollman、Miller&Page，1997；他们给出了多维空间模型的一个计算版本，在那个版本中，候选人在各种行为假设下会移动到中心位置。

12. 提议者可能不得不提议41。

13. McCarty&Meirowitz，2014.他们对这个模型以及应用于政治学领域的其他博弈论模型进行了更加深入的分析。

14. 关于拥有否决权的投票者的影响的进一步的分析，请参见：Tsebelis, 2002。

15. 有一项关于洛杉矶房价的研究估计，通勤成本大约为每小时28美元；请参见：Bajari&Kahn, 2008。

16. 这里的计算过程如下。初始收入等于价格乘以数量，我们可以将它写成$p \times q$。在稀疏的市场中，价格下跌后，收入下降10%，销售额增加8%，因此收入等于$0.9p \times 1.08q = 0.972p \times q$。而对于拥挤的市场，价格下跌后的收入等于$0.9p \times 1.33q = 1.197p \times q$。

21 博弈论模型

1. 正式模型以及相关证明如下。令E_i表示博弈参与者i的努力水平。如果胜出，那么博弈参与者i获得的收益等于M-E_i，否则为-E_i。假设博弈参与者胜出的概率等于他在总努力水平中所占的比例，即：

$$P(iwins)=\frac{E_i}{\sum_{j=1}^{N}E_j}$$

为了求解这个模型的唯一的（对称）纳什均衡，我们考虑这个问题：假设所有其他博弈参与者都选择相同的努力水平E^*，博弈参与者i的努力水平是什么。博弈参与者i的收益等于：

$$\frac{E_i\times M}{E_i+(N-1)E^*}-E_i$$

这个函数的一阶导数等于：

$$\frac{(N-1)E^*\times M}{E_i^2+2(N-1)E^*E_i+(N-1)^2(E^*)^2}-1$$

为了求出最大值，我们令一阶导数等于零，于是可以得出$(N-1)E^*\times M=E_i^2+2(N-1)E^*E_i+(N-1)^2(E^*)^2$。在对称均衡中，$E_i=E^*$。代入上式就可以得出结果。而为了证明一阶导数确实给出了最大值，我们只需检验收益函数的二阶导数是不是为负即可。

2. 欲了解如何梳理网络效应、以及为什么使用快照数据很难得出关于网络效应的结论，请参见：Shalizi&Thomas， 2011。关于聚类行为和属性的实例，请参见：Christakis&Fowler, 2009。

22 合作模型

1. 请参见2005年出版的《科学》杂志创刊125周年纪念特刊。

2. Martin et al., 2008; Biernaskie, 2011。

3. Zaretsky， 1998.在银行的例子中，如果银行能够赚取大量的地理租金，ATM会降低它们利润吗？地理租金指，银行可以利用客户无法承担长途跋涉到另外一个城镇的银行办理业务的成本而获得的额外利润。我要感谢西蒙·威尔基（Simon Wilkie）以及其他一些朋友，他们与我一起讨论了这个例子。

4. 从技术上讲，我们这样说其实是在证明冷酷触发策略是概率性重复囚

徒困境博弈中的均衡策略。其他策略，例如针锋相对策略，在与冷酷触发策略配对时，也可以成为均衡策略。

5. 当另一方博弈参与者采取冷酷触发策略时，选择背叛的博弈参与者在第一个时期内可以获得收益T，但是在未来的任何一次博弈中都只能获得最多为零的收益。如果博弈参与者在每个时间段使用冷酷触发策略，那么在博弈进行的第一期都能获得R的收益。博弈进行两期的概率为P，进行三期的概率为P^2，进行N期的概率为p^{N-1}。因此，期望收益等于：

$$\sum_{t=1}^{\infty} P^{t-1} \times R = \frac{R}{1-P}$$

对$(1+P+P^2+P^3+\cdots) = \frac{1}{1-P}$的证明如下。假设结果为真，然后对等式的两边同时乘以（1-P），于是右侧等于1，左侧等于（1＋P＋P^2＋P^3＋…）-（P＋P^2＋P^3＋…），它也等于1。

6. 我们只需要把保持合作时的那个计算过程重做一遍，并使P下降到$\hat{P}$即可。如果$\hat{P}$不支持合作，那么，继续以概率为P的博弈进行几期后，再以$\hat{P}$的继续概率进行博弈，也不能支持合作。

7. 诺瓦克和卡尔·西格蒙德（Karl　Sigmund）（1998）称之为图像评分。另外也请参见：Bshary&Grutter, 2006。

8. 如果假设它们能够增大鸣鸟的领地，那么这两种形式的攻击都可以映射为背叛。请参见：Godard, 1993。不过，将某些行动映射为具有收益的博弈形式，需要广泛的实地观察。

9. 其中有四对值得重点阐述。当针锋相对策略与冷酷触发策略配对时，两者会永远合作下去，而且每个博弈参与者的平均收益均为3。当欺负好人策略与欺负好人策略配对时，前两轮双方都会背叛，然后永远合作，其平均收益略小于2。当冷酷触发策略与欺负好人策略配对时，冷酷触发策略在第一轮合作、欺负好人策略在第一轮就背叛；在第二轮，双方都背叛；在第三轮和第四轮中，冷酷触发策略背叛和欺负好人策略合作；此后，两者都一直背叛。冷酷触发策略的收益序列可以写成1，2，4，4，然后一直是2，因此平均收益为2^+。欺负好人策略的收益序列可以写成4，3，1，1，然后一直是2，平均收益正好为2。当针锋相对策略与欺负好人策略配对时，针锋相对策略在第一轮合作，而欺负好人策略在第一轮就背叛。第二

轮双方都背叛。在第三轮中，欺负好人策略合作，但是针锋相对策略继续背叛。在第四轮中，欺负好人策略第二次合作，针锋相对策略恢复合作。此后，双方都永远合作。双方的收益均为一个1、一个4、一个2，以及一直为3，因此双方的平均收益均为略小于3。

10. 事实上，始终合作策略和冷酷触发策略都是被针锋相对策略占优的。无论针对什么策略，针锋相对策略至少与始终合作策略或冷酷触发策略一样好（或更好）。针锋相对策略之所以占优永远合作策略，是因为针锋相对策略不会被始终背叛策略和欺负好人策略所“剥削”。针锋相对策略之所以占优冷酷触发策略，是因为它能够与欺负好人策略合作，因为针锋相对策略能够原谅欺负好人策略的背叛，但是冷酷触发策略却不能。

在一个后来成为经典的著名实验中，罗伯特·阿克塞尔罗德要求各个领域的学者参加一项“锦标赛”——提交用于有一定概率继续进行下去的重复囚徒困境博弈中的不同策略。在学者们提交的14个策略中，针锋相对策略表现最佳。然后他又组织了一场，要求大家提交新的策略。这一次有62个人提交了各种策略。针锋相对策略再一次胜出。阿克塞尔罗德将针锋相对策略的成功归结为该策略的如下性质：它会合作，又会惩罚，同时它也是宽容的。冷酷触发策略不宽容，因此冷酷触发策略无法重启与欺负好人策略的合作。请参见：Axelrod，1984。表22-1没有给出针对所有策略的平均收益，因为那将意味着假设每个策略具有相同的可能性。一个种群可能大多数人都采取针锋相对策略；另一个种群可能包括了很大比例的采取欺负好人策略的人；而第三个种群则可能有很多人采取始终合作策略和始终背叛策略。

11. 假设，诱惑收益等于奖励收益的4倍，即，$T=4S$，而且种群中5%的人会合作。这样一来，P必定超过$\frac{20T-R}{20T}$。要维持合作，只需要P超过$\frac{(T-R)}{T}$即可。在$T=4$且$R=3$的情况下，合作的出现要求$P\geq\frac{77}{80}$，而维持合作则要求$P\geq\frac{1}{4}$。接下来给出一般性的证明。假设在种群中，有θ比例的人采取针锋相对策略（或，冷酷触发策略），比例（1-θ）的人采取始终背叛策略。再假设每个人都要与整个种群中的所有其他人博弈。与针锋相对策略配对的针锋相对策略（或冷酷触发策略）在每一期都能获得R的收益，因而其期望

收益为$\frac{R}{(1-P)}$。与始终背叛策略配对的针锋相对策略只能获得-S的回报。而始终背叛策略与始终背叛策略相配对时，双方都只能获得零收益。而且，始终背叛策略与针锋相对策略配对时，始终背叛策略能够获得T的收益。因此，针锋相对策略的平均收益等于$\theta\times\frac{R}{(1-P)}-(1-\theta)\times S$，同时始终背叛策略的平均收益则为$\theta\times T$。这样一来，当且仅当$\frac{R}{1-P}-S\times\frac{(1-\theta)}{\theta}\geqslant T$时，针锋相对策略胜过始终背叛策略。因此，当且仅当下式成立时，针锋相对策略获得更高的收益：

$$P\geqslant\frac{T-R+S\times\frac{(1-\theta)}{\theta}}{T+S\times\frac{(1-\theta)}{\theta}}$$

如果θ很小，那么$\frac{(1-\theta)}{\theta}$将变得非常大，上述条件很难成立。关于合作的出现比合作的维持更加困难的进一步分析，请参见：Boyd, 2006。

12. 这个模型，以及对于它的分析，是我们从诺瓦克的论文中借用的（Nowak, 2006）。诺瓦克证明，重复博弈、声誉和亲属选择也能支持合作。

13. 边缘上的节点会复制表现最好的邻居的行动。根据假设，所有相邻的背叛者获得的收益都等于零。而合作的邻居则可以获得$K\times B$-$D\times C$的收益——当且仅当$\frac{B}{C}\geqslant\frac{D}{K}$时，这一收益大于零。

14. 关于群体选择理论的基本概念，请参见：Wilson, 1975。

15. 对于特劳森和诺瓦克的模型，在这里要简单地介绍一下。将一个由N个人组成的种群划分为M个不同大小的群体。在每个群体中，都应用合作行动模型并为每一个人分配一个表现值。令选中个体i的概率等于i的表现除以所有N个人的表现之和的结果。然后，将这个人的一个克隆体添加到同一个群体中。

如果加入这个克隆体后，该群体的大小超过了阈值$\bar{S}$，那么就以概率（1-q）随机地从这个群体中的某个个体移除，同时以概率q将该群体分成两个群体，每个成员随机地进入其中一个群体中。为了保持群体的总数不变，再随机选择一个现有群体将之移除。当M很大

且群体分裂很少发生（q很小）时，当且仅当$\frac{B}{C} \geqslant 1+\frac{\bar{S}}{M}$成立，合作者的数量就会增加。关于这个模型的更加完整的说明，请参见：Nowak, 2006。

16. 当敏捷管理的倡导者米歇尔·佩鲁索（Michelle Peluso）成了IBM的首席营销官时，她组建了一些相互竞争团队，每一个团队的表现都是其他团队可观察的。然后对表现最好的团队加以奖励（Dan, 2018）。这种类型的敏捷管理实践借鉴了敏捷编程的思想。敏捷编程采用同时推进代码编写、程序测试和与用户互动的方法，取代了标准的按顺序依次推进的瀑布式方法。

17. 还有一种策略是慷慨的针锋相对，即，一开始选择合作，且只在部分时间惩罚背叛行为。在会出现误差的博弈中，这种策略优于针锋相对策略和“赢则坚持，输则改之”策略。请参见：Rand et al., 2009；Wu&Axelrod, 1995。

18. Axelrod&Pienta, 2006.

23 与集体行动有关的问题

1. Hardin, 1968.

2. Diamond, 2005.

3. Ostrom, 2005; Ostrom、Janssen&Anderies, 2007.

4. 为了得出社会最优解，假设每个人在公共物品上花费X。整个种群的总效用等于：

$$N\left[2\sqrt{NX}+\text{收入}-X\right]$$

对X求导，并令得到的导数等于零，可以得到：

$$N\left(\frac{\sqrt{N}}{\sqrt{X}}-1\right)=0$$

得到$X=N$。

为了求解对称纳什均衡，我们假设其他每个人都为公共物品贡献相同的数额，并将这个数额记为A。令Y表示这个人贡献的数额。于是这个人的效用等于：

$$2\sqrt{(N-1)A+Y}+\text{收入}-Y$$

对Y求导，并令得到的导数等于零，我们可以得到$\frac{1}{\sqrt{(N-1)A+Y}}=1$。重新排列各项并对两边同时取平方，得到$Y+(N-1)A=1$。在对称均衡中，每个人贡献相同的数额（$Y=A$），因此有$Y=\frac{1}{N}$。

5. 哲学家会认为这是功利主义。约翰·罗尔斯（John Rawls, 1971）提出了另一种选择：最大最小原则（maxmin principle）。根据这个原则，理想的社会结果就能最大化最不富裕的人的效用。罗尔斯争辩说，我们应该站在“无知之幕”背后去考虑配置问题。“无知之幕”指我们对自己社会中的地位毫不知情。有了“无知之幕”，我们不知道我们是贫是富、有无名声，也不知道自己拥有多大的能力、会不会受到环境的阻碍。罗尔斯要求对最不富裕的人赋予最大的权重，而功利主义则平等地给每个人赋予权重。

6. 用i来给这N个人编号。我们可以把第j个人的效用写成如下形式：

$$\text{Utility}_i(\text{PUBLIC}+\text{PRIVATE})=\sqrt{\text{PUBLIC}}+\text{PRIVATE}_i$$

为了解出对称纳什均衡，我们假设每一个其他人都为公共物品贡献了数量A。

再令Y表示个体j所贡献的数额。用I表示所有人的共同的收入水平。这样一来，个体j的效用等于：

$$(1-\alpha)\left(2\sqrt{(N-1)A+Y}+I-Y\right)+\alpha N\left(2\sqrt{(N-1)A+Y}\right)+\alpha(N-1)(I-A)+(I-Y)$$

对Y求导，并令得到的导数等于零，我们可以得到：

$$\frac{1-\alpha}{\sqrt{(N-1)A+Y}}+\frac{\alpha N}{\sqrt{(N-1)A+Y}}=1$$

重新排列各项，我们有：$(1-\alpha)+\alpha N=\sqrt{Y+(N-1)A}$。在对称平衡中，$Y=A$。从而，遵循$(1-\alpha)+\alpha N=\sqrt{NY}$。等号双边求平方，可以得到$[(1-\alpha)+\alpha N]^2=NY$，这就意味着，$Y=\frac{[(1-\alpha)+\alpha N]^2}{N}$。

7. Cornes & Sandler, 1996.

8. 更现实的模型还要假设非线性拥塞成本（也许是S形的）。这种假设

能够刻画像道路这样的资源的如下特点：前几位其他用户对个人的利益不会有任何影响，但是到了某个时点上，资源会因过于拥挤而归于无用。

9. 使用资源的M个人的总效用等于（B-$\theta\times M$）。对M求导，并令得到的导数等于零，我们可以得到（B-2$M\theta$）＝0，求解可得$M=\frac{B}{2\theta}$。为了求解纳什均衡，我们将从不使用资源的收益设定为零值。于是，人们会使用资源，直到收益等于外部可选项，即：$M=\frac{B}{\theta}$。

10. 需要注意的是：我们将最大收益B设定为等于种群规模N，目的是减少变量的数量。为了求解社会最优结果和纳什均衡，我们首先注意到，总效用等于（N-M）×M＋3N-（N-M）×（N-M），从而得到4（N-M）M。对M求导，得4N-8M＝0。求解可得M＝N/2。于是总效用等于$4(\frac{N}{2})^2=N^2$。为了求解均衡，我们要找到一个M，使两个公园的边际效用相等。这种情况在（N-M）＝3N-3（N-M）时发生。这个式子可以重写为N＝4M。将M和N-M的值代入效用函数，就可以计算出总效用。

11. 为了求解均衡消费，只要设定R^*＝（1-g）（R^*-C^*）并解出R^*即可。

12. Kurlansky, 1998.

13. 为了说明为什么变化没有抵消，可以考虑一个两期模型。第一年的增长率为20%，因此只有96个单位的资源（80×1.2＝96）。第二年的增长率为30%，这导致了98.8＝（96-20）×1.3个单位的资源。如果我们将这两年的增长率互换一下，那么在第一年之后，就会有104个单位的资源，而且在第二年之后会存在100.8＝（104-20）×1.2个单位的资源。

14. Ostrom、Janssen&Anderies, 2007.

15. Craine&Dybzinski, 2013.

16. Ostrom, 2010; Ostrom, 2004.

24　与机制设计有关的模型

1. Ledyard、Porter&Rangel, 1997.

2. 以帕累托有效为例。考虑3个人的以下4种收益情况：

{（3，3，4），（9，0，0），（0，8，1），（2，2，3）}

除了（2，2，2）之外，都是帕累托有效的。但是配置（2，2，3）被（3，3，4）占优。

3. Hurwicz&Schmeidler, 1978.

4. 第三个竞买人可能出价略高于60美元，为了简化分析，假设正好为60美元。

5. 我们在这里给出的证明假设了这些值是均匀分布的，但是这个结果在其他类型的分布下仍然成立。假设其他（N-1）个竞买人的出价都是真实价值乘以$\frac{N-1}{N}$。如果其他竞买人的价值乘以$\frac{N-1}{N}$小于b，那么出价b就将高于另一个竞买人的出价。发生这种情况的概率等于$b\frac{N}{N-1}$。由此可知，大于所有其他（N-1）个出价的概率等于这个值的（N-1）次幂。因此，如果竞买人的真实价值为V，则出价b的预期价值等于该价值减去出价的差（V-b）乘以b为最高出价的概率。这样一来，就可以将预期价值写为：

$$\text{预期价值}=(V-b)\left(b\frac{N}{N-1}\right)^{N-1}$$

为了使上式的值最大化，我们对b求导，并令求得的导数等于零，于是可以得到如下条件：

$$V(N-1)\left(\frac{N}{N-1}\right)^{N-1}b^{N-2}-N\left(\frac{N}{N-1}\right)^{N-1}b^{N-1}=0$$

简化后可以得到这个表达式：V（N-1）-Nb＝0，重写为$b=\frac{V(N-1)}{N}$。为了证明出价最高的竞买人付出的价格等于出价第二高的竞买人的出价的预期价值，我们要先注意到，如果从区间［0，1］上的均匀分布中抽取N个随机的值，那么最高出价的预期价值等于$1\times\frac{N}{N+1}$。第二高的出价的预期价值则等于$1\times\frac{N-1}{N+1}$。因此，出价最高的竞买人的预期价值等于：

$$\left(\frac{N-1}{N}\right)\times1\times\left(\frac{N}{N+1}\right)=1\times\left(\frac{N-1}{N+1}\right)$$

这就等于第二高的竞买人的预期价值。

6. 罗杰·迈尔森（Roger Myerson）是我的博士生导师，他因这个贡献获得了诺贝尔经济学奖。

7. 在全支付拍卖中，当竞买人的价值位于区间［0，1］上时，其最优策略可以表示为：拥有价值V的竞买人出价$V^N \frac{(N-1)}{N}$。因此，如果有三个竞买人，那么价值为$\frac{1}{2}$的竞买人将出价$\frac{1}{8} \times \frac{2}{3} = \frac{1}{12}$。

8. 相关的实验证据，请参见Lucking-Reiley，1999。相关的eBay拍卖实验证据，请参见Morgan&Hossain, 2006。相关的木材拍卖证据，请参见Athey、Levin&Seira, 2011。

9. Ostrovsky、Edelman&Schwarz, 2007.

10. Page, 2012.

25　信号模型

1. 有关信号如何驱动行为和选择的更多例子，请参见：Simler&Hanson, 2018。

2. 一个弱者类型的个体发送信号成本等于MC。假设所有S个强者类型的个体都发出了信号，那么收益等于$\frac{B}{S+1}$。因此，如果$MC \geqslant \frac{B}{S+1}$，那么弱者类型个体就不会发送信号。如果他们源于分离的收益$\frac{B}{S} - cM$，超过了所有N个个体分担信号时的$\frac{B}{N}$，强者类型更偏好发送信号。从先前的计算可知，使弱者类型的个体不发送信号的最小信号等于$M = \frac{B}{(S+1)C}$。因此，如果我们$\hat{M} = \frac{B}{SC}M$，那么弱者类型将不会发出信号。而要让强者类型更偏好发送信号$\hat{M}$，我们必须保证$\frac{B}{S} - c\frac{B}{SC} \geqslant \frac{B}{N}$。上式两边同时除以$B$并乘以$C$，就可以得到$\frac{C}{S} - \frac{c}{S} \geqslant \frac{C}{N}$。这个式子可以简化为（$C-c$）$N \geq CS$，可以改写为$C$（$N-S$）$\geq cN$。

3. 经济学家迈克尔·斯宾塞（Michael Spence）提出了这种可能性。斯宾

塞因提出了一个假设教育没有功能性价值的就业市场模型（Spence, 1973）获得了诺贝尔经济学奖。

4. 尾羽过大就会影响孔雀的适合度。请参见：Zahavi, 1975。

5. Bird&Smith, 2005.

6. Smith&Bird, 2003.

26 学习模型

1. 关于学习的心理学研究，涵盖的背景要比我们在这里所讨论的更加广泛。一个人可以通过学习，了解一个事实，例如阿肯色州的首府在哪里。一个人也可以通过学习，获得某种隐性知识，例如怎样烤面包、修理发动机或编写计算机程序。一个人还可以通过学习，掌握某种知识体系，例如有机化学。

2. Thorndike, 1911.

3. Rescorla&Wagner, 1972.

4. 我描述的这个模型是以下述论著中模型为基础的：Rescorla&Wagner, 1972; Herrnstein, 1970; Bush&Mosteller, 1955; Cyert&March, 1963; Bendor、Diermeie&Ting, 2003; Epstein, 2014。

5. 对参数γ的选择，必须保证备选方案的权重为正。如果γ超过了最高可能渴望水平与任何备选方案的最小奖励之间差异的倒数，就可以做到这一点。

6. Bendor&Swistak, 1997.

7. 如果我们构建出一个有限种群的复制者动态模型，随后的种群是随机选择的，那么就可能无法重新复制出最优备选方案。如果真的是这样，那么复制者动态模型将找不到最优选择，因为它无法将备选方案重新引入到种群中。

8. Fudenberg&Levine, 1998; Camerer, 2003.

9. 这个博弈还有一个混合策略均衡，其中2/3选择经济型汽车，1/3选择一个油老虎车。但是，根据我们的学习规则，这种均衡是不稳定的，因此忽略它。

10. 更正式的表示是，P（经济型汽车，1）＝0.5，P（油老虎车，1）

＝0.5，收益（经济型汽车，1）＝1.5，收益（油老虎车，1）＝2，平均收益＝1.75。应用复制者方程，我们可以得出P（经济型汽车，2）＝$0.5\times\frac{1.5}{1.75}=0.43$，以及$P$（油老虎车，2）＝$0.5\times\frac{2.}{1.75}$＝0.57。

11. Frank, 1985.

12. Waltz, 1979; Powell, 1991.

13. Vriend，2000.这篇论文分析了类似的收益结构，并将它解释为生产相同产品并同时选择数量的企业之间的竞争。经济学家将这种模型称为古诺竞争模型（Cournot competition model）。

14. 罗斯-伊雷夫学习模型是通过以下公式更新第t期中的备选方案k的权重W（k，t）的：W（k，t+1）＝（1-r）×W（k，t）+Δ（k，t，e）。参数r是一个新近度参数（recency parameter）。如果行动k被选中，那么Δ（k，t，e）＝（1-e）×收益（k，t）；如果行动k没有被选中，那么Δ（k，t，e）＝e×收益（k，t）。参数e（实验参数）决定了未被选中的备选方案的权重。

15. Camerer&Ho, 1999.

16. 这个分析紧密追随了贝德纳和佩奇（2007，2018）的思路，并借鉴了格雷夫（2006）的模型。贝德纳和佩奇强调了初始行动对于均衡的涌现的重要性，而格雷夫则更侧重于信念的作用。也请参见：Gilboa&Schmiedler，1995.他们提出了基于案例的决策理论。另外，关于身份在经济选择中的作用，请参见：Akerlof&Kranton, 2010.

17. 为了证明复制者动态的结果，请注意当且仅当以下情况成立时，按文化规范采取行动才会有更高的收益：

$$(1-B)\times200+B\times220>(1-B)\times180+300B$$

重新排列各项，得20（1-B）＞80B。结果得证。

27 多臂老虎机问题

1. 关于这类问题与经济现象的相关性，请参见：Bergemann&Valimaki, 2008。

2. Hills et al., 2015.

3. 欲了解多臂老虎机问题的更多技术细节，以及对各种启发式的详细比较，请参见：Scott, 2010。

4. 吉廷斯和琼斯（1972）率先描述了这些最优规则。可以将吉廷斯指数重新表述为一个贝尔曼方程（Bellman equation）。贝尔曼方程适用于任何通过一系列选择完成、且每个选择都会产生收益的问题。贝尔曼方程依赖于价值函数的构造。该价值函数的值等于收益序列的总和（其中，未来的收益要根据当前利率进行贴现）。

5. Roberts, 2004.

6. 请参见杰克·鲍尔斯（Jake Bowers）等人（2017）对美国农业部农场服务机构小额贷款项目的实验分析。

7. 数据源于2012年的《华盛顿邮报》，以及Dann, 2016。

8. 如果他们是在位领导人，那么阿尔·戈尔（Al Gore）、乔治·H. W. 布什（George H. W. Bush）和希拉里·克林顿将会因经济繁荣而获得更大声誉。关于较早期的政治选情与经济景气的关系的分析，请参见：Markus, 1988。关于更晚近的证据，请参见：Fair, 2012。关于候选人与政党效应大小的相关证据，请参见：Campbell、Dettrey&Yin, 2010。

28 崎岖景观模型

1. 关于多样性的价值的更加深入的阐述，请参见《多样性红利》一书。

2. 对NK模型更加全面的介绍，请参见：Kauffman, 1993。

3. 当$N=20$，$K=19$时，我们可以推导出局部峰值和全局峰值的期望值。每个属性的贡献均匀分布在区间[0，1]上，其均值为$\frac{1}{2}$、方差为$\frac{1}{12}$。这个结果只要通过观察均匀分布就可以证明。每个备选方案的价值等于20个属性的贡献的平均值。因此，根据中心极限定理，这些值服从均值为$\frac{1}{2}$、方差为$\frac{1}{12N}$的正态分布。从而，在$N=20$的情况下，每一个标准偏差的大小为$0.0645=\sqrt{\frac{1}{240}}$。然后我们可以估计出局部峰值的平均值：0.609。我们可以认为，从

这个分布中随机抽取21次当中最好的那一次就是一个局部峰值。因此，它的期望值等于近似地从正态分布中抽取出来的、使$\frac{21}{22}$的备选方案都有较低的值的一个值。它将会比平均值略低两个标准偏差。使用正态分布计算器，可以算出期望平均值等于0.609。为了估计出全局峰值的期望值0.759，我们注意到，全局峰值具有所有2^{20}个备选方案的最高值。每个备选方案可以被视为是从该分布中抽出的一个值，因此其期望值等于近似地从正态分布中抽取出来的、使$\frac{2^{20}}{2^{20}+1}$的备选方案都有较低的值的一个值。利用正态分布计算器，可以算出该期望平均值等于0.759。全局峰值的值，应该比100万次抽取的最高值还更高。

4. Wright, 2001.研究认为这些整合重组有助于人类以及人类社会的出现，也有利于这个时代的技术和科学进步。

5. Kaufman, 1993; Miller&Page, 2007.他们把复杂性描述为“有趣的中间状态”。

6. 在美国，如果申请专利后获批的时间超过了3年，那么专利保护期为从专利发布日起计算17年。

7. Boldrin&Levine, 2010.

结语　像芒格一样智慧地思考——多模型思维的实际应用

1. 根据正文所述的概率，确切的统计均衡是：无痛状态为70.7%，阿片状态为19.5%，上瘾状态为9.8%。而对于第一种情况，相应的百分比则分别为76.3%、21.5%和2.2%。

2. Wakeland、Nielsen&Geissert, 2015.

3. 我要感谢艾比·雅各布斯（Abbie Jacobs），她帮助我得到了这个结果。她有所需要的模型和数据。

4. Wilkinson&Pickett, 2009.

5. 这是杜尔劳夫在芝加哥大学贝克尔·弗里德曼研究所（Becker Friedman Institute）主办的一次会议上的评论。会议的主题是“了解不平等及其应对措施”，于2015年11月6日举行。

6. Goldin&Katz, 2008; Acemoglu&Autor, 2011; Murphy&Topel, 2016.

7. Mas-Colell、Whinston&Green, 1995.

8. Kaplan&Rauh, 2013a; Jones&Kim, 2018.他们的模型将能力作为企业家创意的可扩展性的代理变量。另请参见：Frank, 1996；该文是关于不同职业中的不平等是如何产生的早期研究。另请参见：Xie、Killewald&Near, 2016。

9. Ormerod, 2012.

10. Ormerod，2012.该文详细描述了日益增强的连通性如何导致了不平等。

11. Cancian&Reed, 1999; Schwartz&Mare, 2005.

12. Greenwood et al., 2014.

13. 估计的结果是0.34而不是0.43。请参见：Greenwood et al., 2014。基尼系数可以衡量收入分布与平均分布之间的距离。令$S(P)$表示收入分布中最底部P%的人口所占的总份额（例如，如果位于收入最底部30%的人口获得了总收入的2%，那么$S(30)=2$，计算出来的基尼系数为：

$$\text{基尼系数}=\frac{2}{99}\times\sum_{P=1}^{100}\left[\frac{P}{100}-S(P)\right]$$

如果收入是均匀分布的，即$S(P)=\frac{P}{100}$，那么基尼系数＝0。如果所有的收入都归于最顶部的1%，那么对于所有$P<100$，都有$S(P)=0$，且$S(100)=1$，因此基尼系数＝1。

14. 计算可以按如下过程进行。儿童属于四个类别的概率为（0.6，0.25，0.1，0.05）。这也就是说，高收入者的下一代中60%的人有高收入，中高收入者的下一代中20%的人有高收入，中低收入者的下一代中15%的人有高收入，而低收入者的下一代中只有5%的人有高收入。高收入者的孙子女中有高收入的人的百分比则等于$0.6\times0.6+0.25\times0.2+0.1\times0.15+0.05\times0.05=0.4275$，低收入者的孙子女中有高收入的人的百分比则等于$0.6\times0.05+0.25\times0.1+0.1\times0.15+0.05\times0.7=0.105$。

15. Pfeffer&Killewald, 2017.

16. Kaplan&Rao, 2013b.

参考文献

Acemoglu, Daron, and David Autor. 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics,* 4: 1043–1171. Amsterdam: Elsevier-North Holland.

Acemoglu, Daron, and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Cambridge, MA: Harvard University Press.

Adler, Mortimer Jerome. 1970. *The Time of Our Lives: The Ethics of Common Sense.* New York: Holt, Rinehart and Winston.

Akerlof, G., and R. Kranton. 2010. *Identity Economics*. Princeton, NJ: Princeton University Press.

Albert, Rika, Istvan Albert, and Gary L. Nakarado. 2004. “Structural Vulnerability of the North American Power Grid.” *Physical Review E* 69: 025103.

Allesina, Stefano, and Mercedes Pascual. 2009. “Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions?” *PLOS: Computational Biology* 9, no. 4.

Allison, Graham. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. New York: Little, Brown.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. 2013. “The World Top Incomes Database.” https://www.inet.ox.ac.uk/projects/view/149.

Anderson, Chris. 2008a. “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.” *Wired* 16, no. 7.

Anderson, Chris. 2008b. *The Long Tail: Why the Future of Business Is Selling Less of More.*New York: Hachette.

Anderson, Phillip. 1972. “More Is Different.” *Science* 177, no. 4047: 393–396.

Arrow, Kenneth. 1963. *Social Choice and Individual Values*. New Haven, CT: Yale University Press.

Arthur, W. B. 1994. “Inductive Reasoning and Bounded Rationality (The El Farol Problem).” *American Economic Review Papers and Proceedings* 84: 406–411.

Arthur, W. B. 2011. *The Nature of Technology: What It Is and How It Evolves.* New York: Free Press.

Ashenfelter, Orley. 2010. “Predicting the Quality and Prices of Bordeaux Wine.” *Journal of Wine Economics* 5, no. 1: 40–52.

Athey, Susan, Jonathan Levin, and Enrique Seira. 2011. “Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions.” *Quarterly Journal of Economics* 126, no. 1: 207–257.

Austin, David. 2008. “Percolation: Slipping Through the Cracks.” American Mathematical Society. www.ams.org/publicoutreach/feature-column/fcarc-percolation.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Axelrod, David, Robert Axelrod, and Kenneth J. Pienta. 2006. “Evolution of Cooperation Among Tumor Cells.” *Proceedings of the National Academy of Sciences* 103, no. 36: 13474–13479.

Axtell, Robert L. 2001. “Zipf Distribution of U. S. Firm Sizes.” *Science* 293: 1818–1820.

Bajari, Patrick, and Matthew E. Kahn. 2008. “Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl.” In *Hedonic Methods in Housing Markets,* 129–155. New York: Springer.

Bak, Per. 1996. *How Nature Works: The Science of Self-Organized Criticality.* New York: Springer.

Baldwin, Carliss Y., and Kim B. Clark. 2000. *Design Rules. Vol. 1, The Power of Modularity*.Cambridge, MA: MIT Press.

Ball, Eric, and Joseph LiPuma. 2012. *Unlocking the Ivory Tower: How*

Management Research Can Transform Your Business. Palo Alto, CA: Kauffman Fellow Press.

Banzhaf, John F. 1965. “Weighted Voting Doesn't Work: A Mathematical Analysis.” *Rutgers Law Review* 19, no. 2: 317–343.

Barber, Gerald M. 1997. “Sequencing Highway Network Improvements: A Case Study of South Sulawesi.” *Economic Geography* 53, no. 1: 55–69.

Bass, Frank. 1969. “A New Product Growth Model for Consumer Durables.” *Management Science* 15, no. 5: 215–227.

Baxter, G. William. 2009. “The Dynamics of Foraging Ants.” Paper presented at the annual meeting of the American Physical Society, March 16–20, abstract H40.00011.

Bednar, Jenna. 2007. “Credit Assignment and Federal Encroachment.” *Supreme Court Economic Review* 15: 285–308.

Bednar, Jenna. 2008. *The Robust Federation: Principle of Design.* Cambridge: Cambridge University Press.

Bednar, Jenna, Aaron Bramson, Andrea Jones-Rooy, and Scott E. Page. 2010. “Emergent Cultural Signatures and Persistent Diversity: A Model of Conformity and Consistency.”*Rationality and Society* 22, no. 4: 407–444.

Bednar, Jenna, and Scott E. Page. 2007. “Can Game(s) Theory Explain Culture? The Emergence of Cultural Behavior Within Multiple Games.” *Rationality and Society* 19, no. 1: 65–97.

Bednar, Jenna, and Scott E. Page. 2018. “When Order Affects Performance: Culture, Behavioral Spillovers and Institutional Path Dependence.” *American Political Science Review* 112, no. 1: 82–98.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2018 “Who Becomes an Inventor in America? The Importance of Exposure to Innovation: Executive Summary.” www.equality-of-opportunity.org.

Bendor, Jonathan, Daniel Diermeier, and Michael Ting. 2003. “A Behavioral Model of Turnout.” *American Political Science Review*

97, no. 2: 261–280.

Bendor, Jonathan, and Piotr Swistak. 1997. “The Evolutionary Stability of Cooperation.”*American Political Science Review* 91: 290–307.

Bendor, Jonathan, and Scott E. Page. 2018. “A Model of Team Problem Solving.”Unpublished manuscript.

Berg, Nathan, and Gerd Gigerenzer. 2010. “As-If Behavioral Economics: Neoclassical Economics in Disguise?” *History of Economic Ideas* 18, no. 1: 133–166.

Bergemann, Dirk, and Juuso Välimäki. 2008. “Bandit Problems.” In *The New Palgrave Dictionary of Economics,* 2nd ed., ed. Steven N. Durlauf and Lawrence E. Blume. London: Palgrave Macmillan.

Berlekamp, Elwyn R., John H. Conway, and Richard K. Guy. 1982. “What Is Life?” In *Winning Ways for Your Mathematical Plays*. Vol. 2, *Games in Particular*. London: Academic Press.

Bertrand, Marianne, and Sendhil Mullainathan. 2001. “Are CEOs Rewarded for Luck? The Ones Without Principles Are.” *Quarterly Journal of Economics* 116: 901–932.

Bickel, P. J., E. A. Hammel, and J. W. O'Connell. 1974. “Sex Bias in Graduate Admissions: Data from Berkeley.” *Science* 187 (4175): 398–404.

Biernaskie, Jay, M. 2011. “Evidence for Competition and Cooperation Among Climbing Plants.” *Proceedings of the Royal Society B* 278: 1989–1996.

Bird, Rebecca, and Eric Smith. 2004. “Signaling Theory, Strategic Interaction, and Symbolic Capital.” *Current Anthropology* 46, no. 2: 222–248.

Boldrin, Michele, and David Levine. 2010. *Against Intellectual Monopoly*. Cambridge: Cambridge University Press.

Borges, Jorge Luis. 1974. *A Universal History of Infamy*. Trans. Norman Thomas de Giovanni. London: Penguin.

Bowers, Jake, Nathaniel Higgins, Dean Karlan, Sarah Tulman, and Jonathan Zinman. 2017."Challenges to Replication and Iteration in Field Experiments: Evidence from Two Direct Mail Shots." *American Economic Review Papers & Proceedings* 107, no. 5: 1–3.

Bowles, Samuel, and Herbert Gintis. 2002. "The Inheritance of Inequality." *Journal of Economic Perspectives* 16, no. 3: 3–30.

Box, George E. P., and Norman Draper. 1987. *Empirical Model-Building and Response Surfaces*. New York: Wiley.

Boyd, Robert. 2006. "Reciprocity: You Have to Think Different." *Journal of Evolutionary Biology* 19: 1380–1382.

Breiman, Leo. 1996. "Bagging Predictors." *Machine Learning* 24, no. 2: 123–140.

Briggs, Andrew, and Mark Sculpher. 1998. "An Introduction to Markov Modeling for Economic Evaluation." *Pharmaco Economics* 13, no. 4: 397–409.

Brock, William, and Steven Durlauf. 2001. "Discrete Choice with Social Interactions."*Review of Economic Studies* 68: 235–260.

Broido, A. D., and A. Clauset. 2018. "Scale-Free Networks Are Rare." Working paper.

Bshary, R., and A. S. Grutter. 2006. "Image Scoring and Cooperation in a Cleaner Fish Mutualism." *Nature* 441, no. 7096: 975–978.

Burt, Ronald. 1995. *Structural Holes: The Social Structure of Competition.* Cambridge, MA: Harvard University Press.

Bush, Robert, and Frederick Mosteller. 1954. *Stochastic Models for Learning.* New York: John Wiley and Sons.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.*Princeton, NJ: Princeton University Press.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics* 112, no. 2: 407–441.

Camerer, Colin, and Tek Ho. 1999. “Experience-Weighted Attraction Learning in Normal Form Games.” *Econometrica* 67, no. 4: 827–874.

Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. “Neuroeconomics: How Neuroscience Can Inform Economics.” *Journal of Economic Literature* 43: 9–64.

Campbell, Donald T. 1976. “Assessing the Impact of Planned Social Change.” Public Affairs Center, Dartmouth College.

Campbell, James E., Bryan J. Dettrey, and Hongxing Yin. 2010. “The Theory of Conditional Retrospective Voting: Does the Presidential Record Matter Less in Open-Seat Elections?” *Journal of Politics* 72, no. 4: 1083–1095.

Cancian, Maria, and Deborah Reed. 1999. “The Impact of Wives' Earnings on Income Inequality: Issues and Estimates.” *Demography* 36, no. 2: 173–184.

Carvalho, Vasco, and Xavier Gabaix. 2013 “The Great Diversification and Its Undoing,”*American Economic Review* 103, no. 5: 1697–1727.

Castellano, Claudio, Santo Fortunato, and Vittorio Loreto. 2009. “Statistical Physics of Social Dynamics.” *Review of Modern Physics* 81: 591–646.

Cederman, Lars Erik. 2003. “Modeling the Size of Wars: From Billiard Balls to Sandpiles.”*American Political Science Review* 97: 135–150.

Centola, Damon, and Michael Macy. 2007. “Complex Contagions and the Weakness of Long Ties.” *American Journal of Sociology* 113: 702–734.

Chance, Donald. 2009. “What Are the Odds? Another Look at DiMaggio's Streak.” *Chance* 22, no. 2: 33–42.

Christakis, N. A., and J. Fowler. 2009. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives.* New York: Little, Brown.

Churchland, Patricia, and Terry J. Sejnowski. 1992. *The Computational*

Brain. Cambridge, MA: MIT Press.

Chwe, Michael. 2013. *Jane Austen: Game Theorist.* Princeton, NJ: Princeton University Press.

Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory.” *Quarterly Journal of Economics* 115, no. 1: 147–180.

Clark, Gregory. 2014. *The Son Also Rises: Surnames and the History of Social Mobility.*Princeton, NJ: Princeton University Press.

Clark, William, Matt Golder, and Sona Nadenicheck Golder. 2008. *Principles of Comparative Politics.*Washington, DC: Congressional Quarterly Press.

Clauset, Aaron, M. Young, and K. S. Gleditsch. 2007. “On the Frequency of Severe Terrorist Attacks.” *Journal of Conflict Resolution* 51, no. 1: 58–88.

Cohen, Tyler. 2013. *Average Is Over: Powering America Beyond the Age of the Great Stagnation.* New York: Dutton.

Cooke, Nancy J., and Margaret L. Hilton, eds. 2014. *Enhancing the Effectiveness of Team Science.* Washington, DC: National Academies Press.

Cornes, Richard, and Todd Sandler. 1996. *The Theory of Externalities, Public Goods, and Club Goods.* 2nd ed. Cambridge: Cambridge University Press.

Craine, Joseph, and Ray Dybzinski. 2013. “Mechanisms of Plant Competition for Nutrients, Water and Light.” *Functional Ecology* 27: 833–840.

Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm.*Englewood Cliffs, NJ: Prentice-Hall.

Dan, Avi. 2018. “How Michelle Peluso Is Redefining Marketing at IBM.” *Forbes,* January 18.

Dann, Carrie. 2016. “Pro-Clinton Battleground Ad Spending Outstrips Trump Team by 2.”NBC News, November 4.

Dawes, Robyn. 1979. "The Robust Beauty of Improper Linear Models in Decision Making."*American Psychologist* 34: 571–582.

de Marchi, Scott. 2005. *Computational and Mathematical Modeling in the Social Sciences*.Cambridge: Cambridge University Press.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. "Optimal Versus Naive Diversification: How Inefficient Is the $\frac{1}{N}$ Portfolio Strategy?"*Review of Financial Studies* 22, no. 5: 1915–1953.

Dennett, Daniel C. 1991. *Consciousness Explained*. Boston: Back Bay Books.

Dennett, Daniel C. 1994. *Darwin's Dangerous Idea: Evolution and the Meanings of Life.*New York: Simon & Schuster.

Denrell, Jerker, and Chengwei Liu. 2012. "Top Performers Are Not the Most Impressive When Extreme Performance Indicates Unreliability." *Proceedings of the National Academy of Sciences* 109, no. 24: 9331–9336.

Diamond, Jared. 2005. *Collapse: How Societies Choose to Fail or Succeed*. New York: Viking Penguin.

Dodds, Peter, Robby Muhamad, and DuncanWatts. 2003. "An Experimental Study of Search in Global Social Networks." *Science* 301: 827–829.

Downing, John A., et al. 2006. "The Global Abundance and Size Distribution of Lakes, Ponds, and Impoundments." *Limnology and Oceanography* 51, no. 5: 2388–2397.

Dragulescu, Adrian, and Victor M. Yakovenko. 2001. "Exponential and Power-Law Probability Distributions of Wealth and Income in the United Kingdom and the United States." *Physica A* 299: 213–221.

Drucker, Peter. 1969. *The Age of Discontinuity: Guidelines to Our Changing Society*. New York: Harper and Row.

Dubos, Jean. 1987. *The White Plague: Tuberculosis, Man and Society*. New Brunswick, NJ: Rutgers University Press.

Dunne, Anthony. 1999. *Hertzian Tales: Electronic Products, Aesthetic Experience and Critical Design.* London: Royal College of Art.

Dyson, Freeman. 2004. “A Meeting with Enrico Fermi.” *Nature* 427: 297.

Easley, David, and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly ConnectedWorld.* Cambridge: Cambridge University Press.

Easley, David, Marcos Lopez de Prado, and Maureen O'Hara. 2012. “Flow Toxicity and Liquidity in a High Frequency World.” *Review of Financial Studies* 24, no. 5: 1457–1493.

Easterly, William, and Stanley Fischer. 1995. “The Soviet Economic Decline.” *World Bank Economic Review* 9, no. 3: 341–371.

Ebbinghaus, Herman. 1885. *Memory: A Contribution to Experimental Psychology*. Online in *Classics in the History of Psychology*. http://psychclassics.yorku.ca/Ebbinghaus/index.htm.

Ehrenberg, Andrew. 1969. “Towards an Integrated Theory of Consumer Behaviour.”*Journal of the Market Research Society* 11, no. 4: 305–337.

Einstein, Albert. 1934. “On the Method of Theoretical Physics.” *Philosophy of Science* 1, no. 2: 163–169.

Eliot, Matt, Ben Golub, and Matthew Jackson. 2014. “Financial Networks and Contagion.”*American Economic Review* 104, no. 10: 3115–3153.

Eom, Young-Ho, and Hang-Hyun Jo. 2014. “Generalized Friendship Paradox in Complex Networks: The Case of Scientific Collaboration.” *Scientific Reports* 4: 4603.

Epstein, Josh. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press.

Epstein, Joshua. 2008. “Why Model?” *Journal of Artificial Societies and Social Simulation* 11, no. 4: 12.

Epstein, Joshua. 2014. *Agent Zero: Toward Neurocognitive Foundations*

for Generative Social Science. Princeton, NJ: Princeton University Press.

Ericsson, K. A. 1996. “The Acquisition of Expert Performance: An Introduction to Some of the Issues.” In *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games,* ed. K. A. Ericsson, 1–50. Mahwah, NJ: Erlbaum.

Fair, Raymond. 2012. *Predicting Presidential Elections and Other Things.* 2nd ed. Stanford, CA: Stanford University Press.

Farmer, J. Doyne 2018. “Collective Awareness: A Conversation with J. Doyne Farmer.” *The Edge*. https://www.edge.org/conversation/j doyne farmer-collective-awareness.

Feld, Scott L. 1991. “Why Your Friends Have More Friends than You Do.” *American Journal of Sociology* 96, no. 6: 1464–1477.

Flegal, Katherine M., Brian K. Kit, Heather Orpana, and Barry I. Graubard. 2012.“Association of All-Cause Mortality with Overweight and Obesity Using Standard Body Mass Index Categories: A Systematic Review and Meta-analysis.” *Journal of the American Medical Association* 309, no. 1: 71–82.

Flores, Thomas, and Irfan Nooruddin. 2016. *Elections in Hard Times: Building Stronger Democracies in the 21st Century*. Cambridge: Cambridge University Press.

Florida, Richard. 2005. *Cities and the Creative Class*. New York: Routledge.

Foster, Dean, and H. Peyton Young. 2001. “On the Impossibility of Predicting the Behavior of Rational Agents.” *Proceedings of the National Academy of Sciences* 98, no. 22: 12848–12853.

Frank, Kenneth, et al. 2018. “Teacher Networks and Educational Opportunity.” In *Handbook on the Sociology of Education,* ed. Barbara Schneider and Guan Saw. New York: Oxford University Press.

Frank, Robert. 1984. *Choosing the Right Pond*. Oxford: Oxford University Press.

Frank, Robert. 1996. *The Winner-Take-All Society: Why the Few at the Top Get So Much More than the Rest of Us.* New York: Penguin.

Freeman, Richard, and Wei Huang. 2015. “Collaborating with People Like Me: Ethnic Coauthorship Within the U. S.” *Journal of Labor Economics* 33 no. S1: S289-S318.

Fudenberg, Drew, and David Levine. 1998. *Theory of Learning in Games.* Cambridge, MA: MIT Press.

Fudenberg, Drew, and David Levine. 2006. “A Dual-Self Model of Impulse Control.”*American Economic Review* 96: 1449–1476.

Gammill, James F. , Jr., and Terry A. Marsh. 1988. “Trading Activity and Price Behavior in the Stock and Stock Index Futures Markets in October 1987.” *Journal of Economic Perspectives* 2, no. 3: 25–44.

Gawande, Atul. 2009. *The Checklist Manifesto: How to Get Things Right*. New York: Henry Holt.

Geithner, Timothy. 2014. *Stress Test: Reflections on Financial Crises.* New York: Crown.

Gerschenkron, Alexander. 1952. “Economic Backwardness in Historical Perspective.”In *The Progress of Underdeveloped Areas*, ed. B. F. Hoselitz. Chicago: University of Chicago Press.

Gertner, Jon. 2012. *The Idea Factory: Bell Labs and the Great Age of American Innovation.*New York: Penguin.

Gibrat, Robert. 1931. *Les inégalités economique*. Paris: Sirely.

Gigerenzer, Gerd, and Reinhard Selten. 2002. *Bounded Rationality: The Adaptive Toolbox.*Cambridge, MA: MIT Press.

Gigerenzer, Gerd, and Peter Todd. 2000. *Simple Heuristics That Make Us Smart.* New York: Oxford University Press.

Gilboa, Itzhak, and David Schmeidler. 1994. “Case-Based Decision Theory.” *Quarterly Journal of Economics* 110: 605–639.

Gilovich, Thomas, Amos Tversky, and R. Vallone. 1984. “The Hot Hand in Basketball: On the Misperception of Random Sequences.” *Cognitive Psychology* 17, no. 3: 295–314.

Glaeser, Edward, Bruce Sacerdote, and Jose Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics* 111, no. 2: 507–548.

Glantz, Andrew. 2008. "A Tax on Light and Air: Impact of the Window Duty on Tax Administration and Architecture, 1696–1851." *Penn History Review* 15, no. 2: 18–40.

Glasserman, Paul, and H. Peyton Young. 2014. "Contagion in Financial Networks." Office of Financial Research Working Paper.

Godard, Renee. 1993. "Tit for Tat Among Neighboring Hooded Warblers." *Behavioral Ecology and Sociobiology* 33, no. 1: 45–50.

Gode, Dhananjay K., and Shyam Sunder. 1993. "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality." *Journal of Political Economy* 101, no. 1: 119–137.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race Between Education and Technology*. Cambridge, MA: Harvard University Press.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U. S. Standard of Living Since the Civil War*. Princeton, NJ: Princeton University Press.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78, no. 6: 1360–1380.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83, no. 6: 1360–1443.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2014. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review: Papers & Proceedings* 104, no 5: 348-353.

Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge: Cambridge University Press.

Griliches, Zvi. 1957, 1988. "Hybrid Corn: An Exploration of the

Economics of Technological Change." In *Technology, Education and Productivity: Early Papers with Notes to Subsequent Literature*. New York: Basil Blackwell.

Groseclose, Tim, and James Snyder. 1996. "Buying Supermajorities." *American Political Science Review* 90: 303–315.

Grossman, S., and J. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70, no. 3: 393–408.

Groysberg, Boris. 2012. *Chasing Stars: The Myth of Talent and the Portability of Performance.* Princeton, NJ: Princeton University Press.

Guy, Richard. 1983. "Don't Try to Solve These Problems." *American Mathematical Monthly* 90: 35–41.

Haidt, Jonathan. 2006. *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom.* Basic Books. New York: NY.

Haldene, Andrew. 2012. "The Dog and the Frisbee." Speech given at the Federal Reserve Bank of Kansas City's 36th Economic Policy Symposium, Jackson Hole, WY.

Haldene, Andrew. 2014. "The Dappled World." Speech given at the University of Michigan Law School, Ann Arbor, October 23.

Haldane, John B. S. 1928. "On Being the Right Size." Online version available at http://irl.cs.ucla.edu/papers/right-size.html.

Hardin, Garret. 1968. "The Tragedy of the Commons." *Science* 162, no. 3859: 1243–1248.

Harrell, Frank E. 2001. *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer.

Harstad, Ronald M., and Reinhard Selten. 2013. "Bounded Rationality Models: Tasks to Become Intellectually Competitive." *Journal of Economic Literature* 51, no. 2: 496–511.Harte, John. 1988. *Consider a Spherical Cow.* Mill Valley, CA: University Science Books.

Hathaway, Oona. 2001. “Path Dependence in the Law: The Course and Pattern of Change in a Common Law Legal System.” *Iowa Law Review* 86.

Havel, Václav. 1985. *The Power of the Powerless: Citizens Against the State in Central-Eastern Europe.* Ed. John Keane. Armonk, NY: M. E. Sharpe.

Hawking, Stephen, and Leonard Mlodinow. 2011. *The Grand Design.* New York: Bantam.

Hecht, Jeff. 2008. “Prophecy of Economic Collapse ‘Coming True.’” *New Scientist*. November 17.

Herrnstein, Richard J. 1970. “On the Law of Effect.” *Journal of the Experimental Analysis of Behavior* 13: 243–266.

Hills, Thomas, Peter M. Todd, David Lazer, A. David Redish, Iain D. Couzin, and the Cognitive Search Research Group. 2015. “Exploration Versus Exploitation in Space, Mind, and Society.” *Trends in Cognitive Science* 19, no. 1: 46–54.

Hofstadter, Douglas, and Emmanuel Sander. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking.* New York: Basic Books.

Holland, John. 1975. *Adaptation in Natural and Artificial Systems.* Ann Arbor: University of Michigan Press.

Hong, Lu, and Scott E. Page. 2009. “Interpreted and Generated Signals.” *Journal of Economic Theory* 144: 2174–2196.

Hotelling, Harold. 1929. “Stability in Competition.” *Economic Journal* 39, no. 153: 41–57.

Huffaker, Carl Burton. 1958. “Experimental Studies on Predation: Dispersion Factors and Predator-Prey Oscillations.” *Hilgardia* 27, no. 14: 343–383.

Hurwicz, Leo, and David Schmeidler. 1978. “Outcome Functions Which Guarantee the Existence and Pareto Optimality of Nash Equilibria.” *Econometrica* 46: 144–174.

Inman, Mason. 2011. “Sending Out an SOS.” *Nature Climate Change* 1: 180–183.

International Monetary Fund. 2009. *Global Financial Stability Report*.

Jackson, Matthew. 2008. *Social and Economic Networks.* Princeton, NJ: Princeton University Press.

Jackson, Matthew and Asher Wolinsky. 1996. “A Strategic Model of Social and Economic Networks.” *Journal of Economic Theory* 71: 44–74.

Jacob, Francois. 1977. “Evolution and Tinkering.” *Science* 196: 1161–1166.

Jacobs, Jane. 1989. *Revolving Doors: Sex Segregation and Women's Careers*. Stanford, CA: Stanford University Press.

Johnson, James. 2014. “Models Among the Political Theorists.” *American Journal of Political Science* 58, no. 33: 547–560.

Johnson-Laird, Philip. 2009. *How We Reason.* New York: Oxford University Press.

Jones, Benjamin F., Brian Uzzi, and Stefan Wuchty. 2008. “Multi-University Research Teams: Shifting Impact, Geography and Social Stratification in Science.” *Science* 322: 1259–1262.

Jones, Charles, and Jihee Kim. 2018 “A Schumpeterian Model of Top Income Inequality.”*Journal of Political Economy.* Forthcoming.

Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decisions Under Risk.” *Econometrica* 47, no. 2: 263–291.

Kalyvas, Stathis. 1999. “The Decay and Breakdown of Communist One - Party Systems.”*Annual Review of Political Science* 2: 323–343.

Kamin, Leon J. 1969. “Predictability, Surprise, Attention and Conditioning.” In *Punishment and Aversive Behavior,* ed. B. A. Campbell and R. M. Church, 279–296. New York: Appleton-

Century-Crofts.

Kaplan, Steven, and Joshua D. Rauh. 2013a. “Family, Education, and Sources of Wealth Among the Richest Americans, 1982–2012.” *American Economic Review Papers and Proceedings* 103, no. 3: 158–162.

Kaplan, Steven, and Joshua D. Rauh. 2013b. “It's the Market: The Broad - Based Rise in the Return to Top Talent.” *Journal of Economic Perspectives* 27, no. 3: 35–56.

Karlsson, Bengt. 2016. “The Forest of Our Lives: In and Out of Political Ecology.”*Conservation and Society* 14, no. 4: 380–390.

Kauffman, Stuart. 1993. *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford: Oxford University Press.

Kennedy, John F. 1956. *Profiles in Courage*. New York: Harper & Brothers.

Khmelev, Dmitri, and F. J. Tweedie. 2001. “Using Markov Chains for Identification of Writers.” *Literary and Linguistic Computing* 16, no. 4: 299–307.

Kleinberg, Jon, and M. Raghu. 2015. “Team Performance with Test Scores.” Working paper, Cornell University School of Information.

Knox, Grahame. n.d. “Lost at Sea.” *Insight,* http:// insight.typepad.co.uk /lost_at_sea.pdf.

Kollman, Ken, J. Miller, and S. Page. 1992. “Adaptive Parties in Spatial Elections.”*American Political Science Review* 86: 929–937.

Kooti, Farshad, Nathan O. Hodas, and Kristina Lerman. 2014. “Network Weirdness: Exploring the Origins of Network Paradoxes.” Paper presented at the International Conference on Weblogs and Social Media (ICWSM), March.

Kurlansky, Mark. 1998. *Cod: A Biography of the Fish That Changed theWorld.* New York: Penguin.

Kydland, Finn E., and Edward C. Prescott. 1977. “Rules Rather than Discretion: The Inconsistency of Optimal Plans.” *Journal of*

Political Economy 85, no. 3: 473–491.

Lai, T. L., and Herbert Robbins. 1985. "Asymptotically Efficient Adaptive Allocation Rules."*Advances in Applied Mathematics* 6, no. 1: 4–22.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112, no. 2: 443–477.

Lamberson, P. J., and Scott E. Page. 2012a. "The Effect of Feedback Variability on Success in Markets with Positive Feedbacks." *Economics Letters* 114: 259–261.

Lamberson, P. J., and Scott E. Page. 2012b. "Tipping Points." *Quarterly Journal of Political Science* 7, no. 2: 175–208.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy* 74: 132–157.

Landemore, Helene. 2013. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton, NJ: Princeton University Press.

Lango, Allen H., et al. 2010. "Hundreds of Variants Clustered in Genomic Loci and Biological Pathways Affect Human Height." *Nature* 467, no. 7317: 832–838.

Langville, Amy N., and Carl D. Meyer. 2012. *Who's #1?: The Science of Rating and Ranking*. Princeton, NJ: Princeton University Press.

Lave, Charles, and James G. March. 1975. *An Introduction to Models in the Social Sciences.* Lanham, MD: University Press of America.

Ledyard, John, David Porter, and Antonio Rangle. 1997. "Experiments Testing Multiobject Allocation Mechanisms." *Journal of Economics and Management Strategy* 6, no. 3: 639–675.

Ledyard, John, David Porter, and Randii Wessen. 2000. "A Market-Based Mechanism for Allocating Space Shuttle Secondary Payload Priority." *Experimental Economics* 2, no. 3: 173–195.

Levins, Richard. 1966. "The Strategy of Model Building in Population Biology." *American Scientist* 54: 421–431.

Levinthal, Daniel A. 1997. “Adaptation on Rugged Landscapes.” *Management Science* 43: 934–950.

Levinthal, Daniel. 1991. “RandomWalks and Organizational Mortality.” *Administrative Science Quarterly* 36, no. 3: 397–420.

Levitt, Steven, and Stephen Dubner. 2009. *SuperFreakonomics: Global Cooling, Patriotic Prostitutes, and Why Suicide Bombers Should Buy Life Insurance.* New York: William Morrow.

Lewis, Michael. 2014. *Flash Boys: A Wall Street Revolt.* New York: W. W. Norton.

Limpert, Eckhard, Werner A. Stahel, and Markus Abbt. 2001. “Log-normal Distributions Across the Sciences: Keys and Clues.” *BioScience* 51, no. 5: 341–352.

Little, Daniel. 1998. *Microfoundations, Method, and Causation: On the Philosophy of the Social Sciences.* Piscataway, NJ: Transaction Publishers.

Lo, Andrew W., and A. Craig MacKinlay. 2007. *A Non-Random Walk Down Wall Street.*Princeton, NJ: Princeton University Press.

Lo, Andrew W. 2012. “Reading About the Financial Crisis: A Twenty-One-Book Review.”*Journal of Economic Literature* 50, no. 1: 151–178.

Lucas, Robert. 1976. “Econometric Policy Evaluation: A Critique.” In *The Phillips Curve and Labor Markets,* ed. K. Brunner and A. Meltzer, 19–46. Carnegie-Rochester Conference Series on Public Policy 1. New York: Elsevier.

Lucking-Reiley, David. 1999. “Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet.” *American Economic Review* 89, no. 5: 1063–1080.

MacKenzie, Debora. 2012. “Boom and Doom: Revisiting Prophecies of Collapse.” *New Scientist*, January.

Mannes, Albert E., Jack B. Soll, and Richard P. Larrick. 2014. “The Wisdom of Select Crowds.” *Journal of Personality and Social Psychology* 107: 276–299.

Markowitz, Harold M. 1952. “Portfolio Selection.” *Journal of Finance* 7, no. 1: 77–91.

Markus, Greg B. 1988. “The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis.” *American Journal of Political Science* 32: 137–154.

Martin, Andrew D., and Kevin M. Quinn. 2002. “Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999.” *Political Analysis* 10: 134–153.

Martin, Francis, et al. 2008. “The Genome of *Laccaria bicolor* Provides Insights into Mycorrhizal Symbiosis.” *Nature* 452: 88–92.

Martinez Peria, Maria Soledad, Giovanni Majnoni, Matthew T. Jones, and Winfrid Blaschke.2001. “Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences.” IMF Working Paper no. 01/88.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1994. *Microeconomic Theory*. New York: Oxford University Press.

Mauboussin, Michael. 2012. *The Success Equation: Untangling Skill and Luck in Business*.Cambridge, MA: Harvard University Press.

May, Robert M., Simon A. Levin, and George Sugihara. 2008. “Ecology for Bankers.”*Nature* 451: 893–895.

McCarty, Nolan. 2011. “Measuring Legislative Preferences.” In *Oxford Handbook of Congress,* ed. Eric Schickler and Frances Lee. New York: Oxford University Press.

McCarty, Nolan, and Adam Meirowitz. 2014. *Political Game Theory: An Introduction*.Cambridge: Cambridge University Press.

McKelvey, Richard. 1979. “General Conditions for Global Intransitivities in Formal Voting Models.” *Econometrica* 47: 1085–1112.

McPhee, William N. 1963. *Formal Theories of Mass Behaviour*. New York: Free Press of Glencoe.

Meadows, D., G. Meadows, J. Randers, and W. W. Behrens III. 1972. *The Limits to Growth.* New York: Universe Books.

Medin, Douglas, Will Bennis, and Michel Chandler. 2010. "The Home-Field Disadvantage."*Perspectives on Psychological Science* 5, no. 6: 708–713.

Merriam, Daniel F., and John C. Davis. 2009. "Using Zipf's Law to Predict Future Earthquakes in Kansas." *Transactions of the Kansas Academy of Science* 112, nos.1&2: 127–129.

Merton, Robert C. 1969. "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case." *Review of Economics and Statistics* 51, no. 3: 247–257.

Merton, Robert K. 1963. "Resistance to the Systematic Study of Multiple Discoveries in Science." *European Journal of Sociology* 4, no. 2: 237–282.

Milgrom, Paul, and John Roberts. 1986. "Pricing and Advertising Signals of Product Quality." *Journal of Political Economy* 94, no. 4: 796–821.

Miller, John H. 1998. "Active Nonlinear Tests (ANTs) of Complex Simulation Models."*Management Science* 44, no. 6: 820–830.

Miller, John H. 2015. *A Crude Look at the Whole*. New York: Basic Books.

Miller, John H., and Scott E. Page. 2004. "The Standing Ovation Problem." *Complexity* 9, no. 5: 8–16.

Miller, John H., and Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.

Miller, Joshua B., and Adam Sanjurjo. 2015. "Surprised by the Gambler's and Hot Hand Fallacies: A Truth in the Law of Small Numbers." IGIER Working Paper no. 552.

Mitchell, Melanie. 1996. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.

Mitchell, Melanie. 2009. *Complexity: A Guided Tour.* Oxford: Oxford University Press.

Mlodinow, Leonard. 2009. *The Drunkard's Walk: How Randomness Rules Our Lives.* New York: Penguin.

Mokyr, Joel. 2002. *The Gifts of Athena: Historical Origins of the Knowledge Economy.*Princeton, NJ: Princeton University Press.

Morgan, John, and Tanjim Hossain. 2006. ". . . Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay." *Advances in Economic Analysis & Policy* 6, no. 2: 3.

Moss-Racusin, Corinne, John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2012. "Science Faculty's Subtle Gender Biases Favor Male Students."*Proceedings of the National Academy of Sciences.* 1647–1649.

Munger, Charles. 1994. "A Lesson on Elementary, Worldly Wisdom as It Relates to Investment Management & Business." University of Southern California Business School.

Murphy, Kevin M., and Robert H. Topel. 2016. "Human Capital Investment, Inequality and Growth." *Journal of Labor Economics* 34: 99–127.

Murray, J. D. 1988. "Mammalian Coat Patterns: How the Leopard Gets Its Spots."*Scientific American* 256: 80–87.

Myerson, Roger B. 1999. "On the Value of Game Theory in Social Science." *Rationality and Society* 4: 62–73.

Myerson, Roger B. 1999. "Nash Equilibrium and the History of Economic Theory." *Journal of Economic Literature* 37, no. 3: 1067–1082.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study."*American Economic Review* 85, no. 5: 1313–1326.

Newman, Mark E. 2005. "Power Laws, Pareto Distributions and Zipf's Law." *Contemporary Physics* 46: 323–351.

Newman, Mark E. 2010. *Networks: An Introduction.* Oxford: Oxford

University Press.

Nowak, Martin. 2006. "Five Rules for the Evolution of Cooperation." *Science* 314, no.5805: 1560–1563.

Nowak, Martin A., and Karl Sigmund. 1998. "Evolution of Indirect Reciprocity by Image Scoring." *Nature* 393: 573–577.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

O'Neil, Cathy 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* New York, NY: Crown.

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349: 6251.

Organization for Economic Co-operation and Development. 1996. *The Knowledge Based Economy*. Paris: OECD.

Ormerod, Paul. 2012. *Positive Linking: How Networks Can Revolutionise the World.*London: Faber and Faber.

Ostrom, Elinor. 2004. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.

Ostrom, Elinor. 2010. "Beyond Markets and States: Polycentric Governance of Complex Economic Systems." *Transnational Corporations Review* 2, no. 2: 1–12.

Ostrom, Elinor, Marco A. Janssen, and John M. Anderies. 2007. "Going Beyond Panaceas." *Proceedings of the National Academy of Sciences* 104: 15176–15178.

Ostrovsky, Michael, Benjamin Edelman, and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review* 97, no. 1: 242–259.

Paarsch, Harry J., and Bruce S. Shearer. 1999. "The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting Industry." *Journal of Human Resources* 34, no. 4: 643–667.

Packer, Craig, and Anne E. Pusey. 1997. “Divided We Fall: Cooperation Among Lions.”*Scientific American,* May, 52–59.

Paczuski, Maya, and Kai Nagel. 1996. “Self-Organized Criticality and 1 / *f* Noise in Traffic.”arXiv: cond-mat/9602011.

Page, Scott E. 1997. “An Appending Efficient Algorithm for Allocating Public Projects with Complementarities,” *Journal of Public Economics* 64, no 3: 291–322.

Page, Scott E. 2001. “Self Organization and Coordination.” *Computational Economics* 18: 25–48.

Page, Scott E. 2006. “Essay: Path Dependence.” *Quarterly Journal of Political Science* 1: 87–115.

Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Teams, Schools, and Societies*. Princeton, NJ: Princeton University Press.

Page, Scott E. 2010a. *Diversity and Complexity*. Princeton, NJ: Princeton University Press.

Page, Scott E. 2010b. “Building a Science of Economics for the Real World.” Presentation to the House Committee on Science and Technology Subcommittee on Investigations and Oversight, July 20.

Page, Scott E. 2012. “A Complexity Perspective on Institutional Design.” *Politics, Philosophy and Economics* 11: 5–25.

Page, Scott E. 2017. *The Diversity Bonus*. Princeton, NJ: Princeton University Press.

Pan, Jessica. 2014. “Gender Segregation in Occupations: The Role of Tipping and Social Interactions.” *Journal of Labor Economics* 33, no. 2: 365–408.

Parrish, Susan Scott. 2017. *The Flood Year 1927: A Cultural History*. Princeton, NJ: Princeton University Press.

Parsa, H. G., John T. Self, David Njite, and Tiffany King. 2005. “Why Restaurants Fail.”*Cornell Hospitality Quarterly* 46, no. 3: 304–322.

Patel, Kayur, Steven Drucker, James Fogarty, Ashish Kapoor, and Desney Tan. 2011.“Using Multiple Models to Understand Data.” *Proceedings of the International Joint Conference on Artificial Intelligence*, 1723-1728.

Peel, L., and A. Clauset. 2014. “Predicting Sports Scoring Dynamics with Restoration and Anti-Persistence.” *Proceedings of the International Conference on Data Mining*.Philadelphia: SIAM.

Pfeffer, Fabian T., and Alexandra Killewald. 2017. “Generations of Advantage: Multigenerational Correlations in Family Wealth.” *Social Forces,* 1–31.

Piantadosi, Steven. 2014. “Zipf'sWord Frequency Law in Natural Language: A Critical Review and Future Directions.” *Psychonomic Bulletin & Review* 21, no. 5: 1112–1130.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

Piketty, Thomas. 2014. *Capital in the 21st Century*. Trans. Arthur Goldhammer. Cambridge, MA: Belknap Press.

Pollack, John. 2014. *Shortcut: How Analogies Reveal Connections, Spark Innovation, and Sell Our Greatest Ideas.* New York: Gotham.

Poole, Keith T., and Howard Rosenthal. 1984. “A Spatial Model for Legislative Roll Call Analysis.” *American Journal of Political Science* 29, no. 2: 357–384.

Porter, David, and Vernon Smith. 2007. “FCC Spectrum Auction Design: A 12-Year Experiment.” *Journal of Law, Economics, and Policy* 3, no. 1: 63–80.

Powell, Robert. 1991. “Absolute and Relative Gains in International Relations Theory.”*American Political Science Review* 85, no. 4: 1303–1320.

Przeworski, Adam, Jose Antonio Cheibub, Michael E. Alvarez, and Fernando Limongi.2000. *Democracy and Development: Political Institutions and Material Well-Being in the World, 1950–1990.* Cambridge: Cambridge University Press.

Raby, Fiona. 2001. *Design Noir: The Secret Life of Electronic Objects*. Basel: Birkhauser.

Ramo, Joshua Cooper. 2016. *The Seventh Sense: Power, Fortune, and Success, in the Age of Networks.* New York: Little, Brown and Company.

Rand, David G., Hisashi Ohtsukia, and Martin A. Nowak. 2009. "Direct Reciprocity with Costly Punishment: Generous Tit-for-Tat Prevails." *Journal of Theoretical Biology* 256, no. 1: 45–57.

Rapoport, Anatol. 1978. "Reality-Simulation: A Feedback Loop." *Sociocybernetics,* 123–141.

Rauch, Jeffrey. 2012. *Hyperbolic Partial Differential Equations and Geometric Optics.*Graduate Studies in Mathematics. Providence, RI: American Mathematical Society.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rescorla, Robert, and Allan Wagner. 1972. "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement." In *Classical Conditioning II,* ed. A. H. Black and W. F. Prokasy, 64–99. New York: Appleton-Century-Crofts.

Reynolds, Noel B., and Arlene Saxonhouse. 1994. *Three Discourses.* Chicago: University of Chicago Press.

Roberts, D. C., and D. L. Turcotte. 1998. "Fractality and Self-Organized Criticality of Wars."*Fractals* 6: 351–357.

Roberts, Seth. 2004. "Self-Experimentation as a Source of New Ideas: Ten Examples About Sleep, Mood, Health, and Weight." *Behavioral and Brain Sciences* 27, no. 2: 227–262

Romer, Paul. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94: 1002–1037.

Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review* 71: 845–858.

Roth, Alvin, and Ido Erev. 1995. "Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economics Behavior* 8: 164–212.

Russakoff, Dale. 2015. *The Prize: Who's in Charge of America's Schools?* Boston: Houghton Mifflin Harcourt.

Rust, Jon. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55, no. 5: 999–1033.

Ryall, Michael D., and Aaron Bramson *Inference and Intervention: Causal Models for Business Analysis.* New York: Routledge.

Salganik, Matthew, Peter Dodds, and Duncan J. Watts. 2006. "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market." *Science* 311: 854–856.

Samuelson, Paul. 1964. "Proof That Properly Anticipated Prices Fluctuate Randomly."*Industrial Management Review* 6: 41–49.

Schiller, Robert. 2004. *Irrational Exuberance.* 2nd ed. Princeton, NJ: Princeton University Press.

Schrodt, Philip. 1998. "Pattern Recognition of International Crises Using Hidden Markov Models." In *Non-Linear Models and Methods in Political Science,* ed. Diana Richards. Ann Arbor: University of Michigan Press.

Schwartz, Christine R., and Robert D. Mare. 2004. "Trends in Educational Assortative Marriage from 1940 to 2003." *Demography* 42, no. 4: 621–646.

Schelling, Thomas. 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton.

Scott, Steven L. 2010. "A Modern Bayesian Look at the Multi-Armed Bandit." *Applied Stochastic Models in Business and Industry* 26: 639–658.

Shalizi, Cosma, and Andrew C. Thomas. 2011. "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies." *Sociological Methods and Research* 40: 211–239.

Shapiro, Thomas, Tatjana Meschede, and Sam Osoro. 2013. “The Roots of theWidening RacialWealth Gap: Explaining the Black-White Economic Divide.” Research and Policy Brief, Institute on Assets and Social Policy, Brandeis University, Waltham, MA.

Shi, Xiaolin, Lada A. Adamic, Belle L. Tseng, and Gavin S. Clarkson. 2009. “The Impact of Boundary Spanning Scholarly Publications and Patents.” *PLoS ONE* 4, no. 8: e6547.

Silver, Nate. 2012. *The Signal and the Noise: Why So Many Predictions Fail— but Some Don't.* New York: Penguin.

Simler, Kevin, and Robin Hanson. 2018. *The Elephant in the Brain: Hidden Motives in Everyday Life*. Oxford: Oxford University Press.

Simmons, Matthew, Lada Adamic, and Eytan Adar. 2011. “Memes Online: Extracted, Subtracted, Injected, and Recollected.” Paper presented at the International Conference on Web and Social Media.

Slaughter, Ann Marie. 2017. *The Chessboard and the Web: Strategies of Connection in a Networked World.* New Haven, CT: Yale University Press.

Smaldino, Paul. 2013. “Measures of Individual Uncertainty for Ecological Models: Variance and Entropy.” *Ecological Modelling* 254: 50–53.

Small, Dana M., Robert J. Zatorre, Alain Dagher, Alan C. Evans, and Marilyn Jones-Gotman. 2001. “Changes in Brain Activity Related to Eating Chocolate: From Pleasure to Aversion.” *Brain* 124, no. 9: 1720–1733.

Smith, Eric, Rebecca Bliege Bird, and D. Bird. 2003. “The Benefits of Costly Signaling: Meriam Turtle Hunters.” *Behavioral Ecology* 14: 116–126.

Smith, Vernon. 2002. “Constructivist and Ecological Rationality.” Nobel Prize lecture.

Sneppen, Kim, Per Bak, Henrik Flyvbjerg, and Mogens Jensen. 1994. “Evolution as a Self-Organized Critical Phenomenon.” *Proceedings*

of the National Academy of Sciences 92: 5209–5213.

Solow, Robert M. 1956. “A Contribution to the Theory of Economic Growth.” *Quarterly Journal of Economics* 70, no. 1: 65–94.

Spence, A. Michael. 1973. “Job Market Signaling.” *Quarterly Journal of Economics* 87, no.3: 355–374.

Squicciarini, Mara, and Nico Voigtländer. 2015. “Human Capital and Industrialization: Evidence from the Age of Enlightenment.” *Quarterly Journal of Economics* 30, no. 4: 1825–1883.

Starfield, Anthony, Karl Smith, and Andrew Bleloch. 1994. *How to Model It: Problem Solving for the Computer Age.* Minneapolis, MN: Burgess International.

Stein, Richard A. 2011. “Superspreaders in Infectious Diseases.” *International Journal of Infectious Diseases* 15, no. 8: e510–e513.

Sterman, John D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. New York: McGraw-Hill.

Sterman, John. 2006. “Learning from Evidence in a ComplexWorld.” *American Journal of Public Health* 96, no. 3: 505–515.

Stiglitz, Joseph. 2013. *The Price of Inequality: How Today's Divided Society Endangers Our Future.* New York: W. W. Norton.

Stock, James H., and Mark W. Watson. 2003. “Has the Business Cycle Changed and Why?” In *National Bureau of Economic Research Macroeconomics Annual 2002*, vol.17, ed. Mark Gertler and Kenneth Rogoff, 159–218. Cambridge, MA: MIT Press.

Stone, Lawrence D., Colleen M. Keller, Thomas M. Kratzke, and Johan P. Strumpfer. 2014.“Search for theWreckage of Air France Flight AF 447.” *Statistical Science* 29, no. 1: 69–80.

Storchmann, Karl. 2011. “Wine Economics: Emergence, Developments, Topics.” *Agrekon* 50, no. 3: 1–28.

Suki, Bela, and Urs Frey. 2017. “A Time Varying Biased RandomWalk Model of Growth: Application to Height from Birth to Childhood.” *Journal of Critical Care* 38: 362–370.

Suroweicki, James. 2006. *The Wisdom of Crowds*. New York: Anchor Press.

Syverson, Chad. 2007. “Prices, Spatial Competition, and Heterogeneous Producers: An Empirical Test.” *Journal of Industrial Economics* 55, no. 2: 197–222.

Taleb, Nassim. 2001. *Fooled by Randomness*. New York: Random House.

Taleb, Nassim. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.

Taleb, Nassim. 2012. *Antifragile: Things That Gain from Disorder*. New York: Random House.

Tassier, Troy. 2013. *The Economics of Epidemiology*. Amsterdam: Springer.

Tetlock, Phillip. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?*Princeton, NJ: Princeton University Press.

Thaler, R. H. 1981. “Some Empirical Evidence on Dynamic Inconsistency.” *Economic Letters* 8, no. 3: 201–207.

Thompson, Derek. 2014. “How You, I, and Everyone Got the Top 1 Percent All Wrong: Unveiling the Real Story Behind the Richest of the Rich.” *Atlantic,* March 30.

Thorndike, Edward L. 1911. *Animal Intelligence.* New York: Macmillan.

Tilly, Charles. 1998. *Durable Inequality*. Berkeley: University of California Press.

Tsebelis, George. 2002. *Veto Players: How Political InstitutionsWork.* Princeton, NJ: Princeton University Press.

Turchin, Peter. 1998. *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants.* Sunderland, MA: Sinauer Associates.

Tweedle, Valerie, and Robert J. Smith. 2012. “A Mathematical Model of Bieber Fever: The Most Infectious Disease of Our Time?” In *Understanding the Dynamics of Emerging and Re-Emerging Infectious*

Diseases Using Mathematical Models, ed. Steady Mushayabasa and Claver P. Bhunu. Cham, Switzerland: Springer.

Ugander, Johan, Brian Karrer, Lars Backstrom, and Cameron Marlow. 2011. "The Anatomy of the Facebook Social Graph." arXiv: 1111.4503.

Updike, John. 1960. "Hub Fans Bid Adieu." *New Yorker,* October 22.

US Bureau of Labor Statistics. 2013. *Consumer Expenditures in 2011*. Report 1042, April. Washington, DC: BLS.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. "Atypical Combinations and Scientific Impact." *Science* 342: 468–471.

Van Noorden, Richard. 2015. "Interdisciplinary Research by the Numbers." *Nature,*September 16.

von Neumann, John, and Morgenstern, Oskar. 1953. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Vriend, Nicolaas J. 2000. "An Illustration of the Essential Difference Between Individual and Social Learning, and Its Consequences for Computational Analyses." *Journal of Economic Dynamics and Control* 24: 1–19.

Wainer, Howard. 2009. *Picturing the UncertainWorld.* Princeton, NJ: Princeton University Press.

Wakeland, W., A. Nielsen, and P. Geissert. 2015. "Dynamic Model of Nonmedical Opioid Use Trajectories and Potential Policy Interventions." *American Journal of Drug and Alcohol Abuse* 41, no. 6: 508–518.

Waltz, Kenneth. 1979. *Theory of International Politics*. New York: McGraw-Hill.

Washington Post. 2012. "Mad Money: TV Ads in the 2012 Presidential Campaign." http://www.washingtonpost.com/wp-srv/special/politics /track-presidential-campaignads-2012.

Watts, Duncan. 2011. *Everything Is Obvious Once You Know the Answer*.

New York: Crown Business.

Watts, Duncan, and Steven Strogatz. 1998. “Collective Dynamics of ‘Small-World’Networks.” *Nature* 393, no. 6684: 440–442.

Weisberg, Michael. 2007. “Three Kinds of Idealization.” *Journal of Philosophy* 104, no. 12: 639–659.

Weisberg, Michael. 2012. *Simulation and Similarity: Using Models to Understand the World.* Oxford: Oxford University Press.

Weisberg, Michael, and Muldoon, Ryan. 2009. “Epistemic Landscapes and the Division of Cognitive Labor.” *Philosophy of Science* 76, no. 2: 225–252.

Weitzman, Martin L. 1979. “Optimal Search for the Best Alternative.” *Econometrica* 77: 641–654.

Weitzman, Martin L. 1998. “Recombinant Growth.” *Quarterly Journal of Economics* 2: 331–361.

Wellman, Michael. 1990. “Fundamental Concepts of Qualitative Probabilistic Networks.”*Artificial Intelligence* 44: 257–303.

Wellman, Michael. 2013. “Head to Head: Does US High-Frequency Trading Need Stricter Regulatory Oversight? (YES).” *International Financial Law Review,* September.

West, Geoffrey. 2017. *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. New York: Penguin.

Whittle, Peter. 1979. “Discussion of Dr Gittins' Paper.” *Journal of the Royal Statistical Society, Series B* 41, no. 2: 148–177.

Whitty, Robin W. 2017. “Some Comments on Multiple Discovery in Mathematics.” *Journal of Humanistic Mathematics* 7, no. 1: 172–188.

Wigner, Eugene. 1960. “The Unreasonable Effectiveness of Mathematics in the Natural Sciences.” *Communications in Pure and Applied Mathematics* 13, no. 1.

Wilkinson, Richard, and Kate Pickett. 2009. *The Spirit Level: Why*

Greater Equality Makes Societies Stronger. London: Bloomsbury.

Wilson, David Sloan. 1975. “A Theory of Group Selection.” *Proceedings of the National Academy of Sciences* 72, no. 1: 143–146.

Wolfram, Stephen. 2001. *A New Kind of Science*. Champaign, IL: Wolfram Media.

Wright, Robert. 2001. *Nonzero: The Logic of Human Destiny*. New York: Vintage.

Wu, Jianzhong, and Robert Axelrod. 1995. “How to Cope with Noise in the Iterated Prisoner's Dilemma.” *Journal of Conflict Resolution* 39, no. 1: 183–189.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. “The Increasing Dominance of Teams in the Production of Knowledge.” *Science* 316, no. 5827: 1036–1039.

Xie, Yu. 2007. “Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science.” *Research in Social Stratification and Mobility* 25: 141–156.

Xie, Yu, Alexandra Killewald, and Christopher Near. 2016. “Between - and Within-Occupation Inequality: The Case of High Status Professions.” *Annals of the American Academy of Political and Social Science* 663, no. 1: 53–79.

Youn, Hyejin, Deborah Strumsky, Luis Bettencourt, and José Lobo. 2015. “Inventions as a Combinatorial Process: Evidence From US Patents.” *Journal of the Royal Society Interfaces* 12: 0272.

Zagorsky, Jay. 2007. “Do You Have to Be Smart to Be Rich? The Impact of IQ onWealth, Income and Financial Distress.” *Intelligence* 35: 489–501.

Zahavi, Amotz. 1974. “Mate Selection: A Selection for a Handicap.” *Journal of Theoretical Biology* 53, no. 1: 205–214.

Zak, Paul, and Stephen Knack. 2001. “Trust and Growth.” *Economic Journal* 111, no.470: 295–321.

Zaretsky, Adam. 1998. “Have Computers Made Us More Productive? A

Puzzle.” *Regional Economist,* Federal Reserve Bank of St. Louis.

Ziliak, Stephen T., and Deirdre N. McCloskey. 2008. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives.* Ann Arbor: University of Michigan Press.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读App玩转指南

湛庐阅读App结构图：



三步玩转湛庐阅读App：



App获取方式：

安卓用户前往各大应用市场、苹果用户前往App Store
直接下载“湛庐阅读”App，与最聪明的人共同进化！

使用App扫一扫功能，
遇见书里书外更大的世界！



湛庐CHEERS

延伸阅读

《多样性红利》

◎　一个人是否聪明不是由智商决定的，而取决于认知工具的多样性！

◎ 《多样性红利》创造性提出多样性视角、启发式、解释和预测模型四个认知工具箱框架。

◎ 广受欢迎的“模型思维课”主讲人斯科特·佩奇经典名作！





《表象与本质》

◎ 闻名世界的认知科学家侯世达凭借独特的智慧与天赋，联合法国心理学家桑德尔，终于向世人展示了这部极具开创性的著作，一解人类认知之谜。

◎ 人类大脑中的每个概念都源于多年来不知不觉中形成的一长串类比，这些类比赋予每个概念生命，我们在一生中不断充实这些概念。大脑无时无刻都在作类比。类比，就是思考之源和思维之火。





《复杂经济学》

◎ 作为“复杂经济学”的创始人，布莱恩·阿瑟在本书中汇集了多年对复杂经济学的研究。其核心思想可以归结为：经济不一定处于均衡状态，演绎推理将被归纳推理所取代。

◎ 这是一本见证“复杂经济学”成长的著作，你将看到“复杂性思维”在经济学领域是如何发展起来并形成一门崭新的学科。

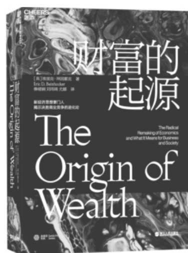
◎ 详细讨论了“爱尔法鲁酒吧问题”和“圣塔菲人工股票市场”两个重要项目。





《财富的起源》

- 物联网时代的《国富论》！牛津大学新经济思想掌门人埃里克·拜因霍克重磅力作！用复杂性重新定义经济、商业和财富，帮你理解复杂时代的复杂经济！
- 《财富的起源》回溯了经济学的发展脉络，用动态、主体、互动、涌现和进化5大视角审视当下经济，阐释商业设计、物理技术、社会技术促进财富增长的进化力量，提出应用复杂经济学的4大应用领域：战略、组织、金融与公共政策，助你利用这部物联网时代的《国富论》决胜商业社会！





(1) 二人均为简·奥斯汀的小说《爱玛》中的人物。——编者注

(2) 重言式，又称永真式。逻辑学名词。如果一个公式，对于它的任一解释，其真值都为真，就称为重言式。经济上称为套套逻辑。——编者注

(3) 乔纳森·海特的著作《象与骑象人》详细讲述了这个比喻。其中文简体字版已由湛庐文化策划，浙江人民出版社出版。——编者注

(4) "爱尔法鲁酒吧问题"最早由经济学家布莱恩·阿瑟提出，他的著作《复杂经济学》《技术的本质》中文简体字版已由湛庐文化策划，浙江人民出版社出版。——编者注

(5) 蒙特·卡罗方法，又称模拟统计方法，是一种以概率统计理论为指导的一类重要的数值计算方法，是使用随机数来解决很多计算问题的方法。——编者注

(6) 丹尼尔·丹尼特是著名的哲学家，他的另一本著作《直觉泵和其他思考工具》中文简体字版已由湛庐文化策划，浙江教育出版社出版。——编者注

(7) 大宗商品指无法区分的产品，如原油、猪肚和2号红小麦等。——编者注

(8) 戈雅，全名弗朗西斯科·何塞·德·戈雅·卢西恩特斯（Francisco José de Goya Lucientes），西班牙浪漫主义画派画家，代表作《裸体的马哈》。——编者注

(9) 卡特尔，由一系列生产类似产品的独立企业所构成的组织，目的是提高该类产品价格并控制其产量。——编者注

CHEERS
湛庐

财富的起源

[英] 埃里克·拜因霍克 著
Eric D. Beinhocker
俸绪娴 刘玮琦 尤娜 译

新经济思想掌门人
揭示决胜商业竞争的进化论

The Origin of Wealth

The Radical Remaking of Economics and What It Means for Business and Society


圣塔菲书系 Santa Fe
Series


浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2019年10月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有•侵权必究

书名：财富的起源

著者： 埃里克•拜因霍克

电子书定价：143.99元

The Origin of Wealth: The Radical Remaking of Economics and What it

Means for Business and Society

Copyright © 2006 by Eric D. Beinhocker

Eric Beinhocker

埃里克•拜因霍克

牛津大学新经济思想研究所掌门人

圣塔菲研究所外聘教授

麦肯锡全球研究所合伙人



牛津大学
新经济思想研究所掌门人

在一次田野调查中，拜因霍克前往了肯尼亚西南地区的一个偏远乡村。在那里，他偶遇了一位马赛族人。那位马赛族人一直在礼貌地询问关于拜因霍克的家乡和家庭的问题，打听他的经济实力。在得知拜因霍克家里一头牛都没有时，马赛族人显露出了遗憾的神情，同时还有一丝疑惑，他搞不明白这样穷困潦倒、家里一头牛都没有的人怎么有钱跑到村子里探访，还拥有相机这种看上去非常昂贵的东西。

后来，在得知拜因霍克的叔叔在马里兰州的农场上养了一大群牛时，马赛族人心里的疑惑“解开了”——很显然，这位客人是仰仗富人叔叔的游手好闲之辈，是靠富人叔叔的财富来旅行和生活的。

这次调查促使拜因霍克想要了解财富是什么，财富是如何被创造出来的，如何才能创造更多的财富等问题。拜因霍克是牛津大学新经济思想研究所的掌门人，牛津大学布拉瓦尼克政府学院公共政策实践教授。而牛津大学新经济思想研究所是一个致力于将前沿的跨学科方法应用于解决金融系统稳定性、创新和增长、经济不平等和环境可持续性等问题的机构，面对经济学领域正在发生的深刻变化，该研究所在开创性研究方面占据着举足轻重的地位。

圣塔菲研究所外聘教授



拜因霍克还是圣塔菲研究所的外聘教授，在《复杂经济学》的作者布莱恩•阿瑟的组织和合作下，他们开展了多项复杂经济学研究项目，在该领域产生了重要影响。

经济学领域正在经历一场变革，由传统的古典经济学研究范式向复杂经济学转变。

在实践和研究中，人们逐渐发现传统经济学中的假设和实践并不符合真实的经济运行状况，比如“完全理性的经济人假设”“均衡理论”等，事实上，经济本身是一个开放的、动态的非线性系统，并不会实现均衡状态，而经济人本身也是不完全理性的，经济本身是由众多个体组成的互动网络，是一种涌现现象，也是一个进化系统。复杂经济学概念将在未来数十年为经济学理论和实践提供支持。圣塔菲研究所正是在经济学领域的范式转移中诞生的专门性研究机构，是世界知名的复杂性科学研究中心。

麦肯锡全球研究所合伙人



在加入牛津大学之前，拜因霍克在麦肯锡公司工作了18年，担任麦肯锡全球研究所合伙人。《财富》杂志称其为“新世纪的商业领袖”。他写作的商业和经济学文章发表在《金融时报》《泰晤士报》《卫报》《大西洋月刊》《新闻周刊》等杂志报纸上。

拜因霍克的文章深入浅出、逻辑严密，用语流畅清晰，获得了读者的诸多好评。在《财富的起源》中，作者从位于菲律宾马尼拉的一处名叫乐土的垃圾场入手，将位于垃圾场上的复杂经济网络向读者娓娓道来：垃圾捡拾者、垃圾贩子、垃圾回收商、制造商以及向垃圾拾荒者及其家人兜售日常用品的小贩，最终引出了经济源自何处，为何一堆垃圾上会诞生如此复杂的经济系统，这些行为、关系、组织和理念如何支撑起一种经济形式等问题。

在作者抽丝剥茧般的叙述中，我们将深入经济学领域看似深不可测的腹地，探索那些趣味十足又具有真正的指导意义的精妙理论和案例。

作者演讲洽谈，请联系

speech@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS 特别制作

献给蒂莉。

自我们用后腿站立以来，

门已经被砸开了五六次。

或许这是活着的最好时机，

当你自认为所知道的一切几乎都是错误的时候。

——汤姆·斯托帕德

赞誉

工业革命以来，人类巨大财富文明的创造过程一直闪耀着亚当·斯密《国富论》的智慧之光。万物互联时代，传统经济理论到了需要改写的关头。埃里克·拜因霍克的著作恰逢其时，堪称物联网时代的《国富论》。

张瑞敏

海尔集团董事局主席、首席执行官

终于有一部讲述经济学思想演进、探究经济学思想前沿的好书！对于痛感传统经济学无法解释数字经济、信息社会和智能未来的人来说，拜因霍克的这部《财富的起源》，堪称与新时代共舞的杰作。

段永朝

苇草智库创始合伙人

经济学曾经和进化论有过一段相互学习、相互激励的时期，但很快就分道扬镳了。分手之后的进化论慢慢沉淀为现代科学的基石，而分手之后的经济学却飘飘在上，看似霞光万道，其实一阵风就能吹走。《财富的起源》回到了进化论和经济学最初相遇的地方，然后尝试让经济学再走一条不一样的道路。这条路当然更为艰辛危险，却能通往正确的方向。

何帆

上海交通大学安泰经济与管理学院经济学教授

本书革命性地展望了复杂经济学的格局，可以预见，以动态、主体、互动、涌现、进化为特征的新的经济学范式，将指导人类更好地理解和把握量子时代的经济发展实践。

陈劲

清华大学经济管理学院教授、知名创新管理学家

《财富的起源》一书不仅用动态、进化等五种视角重新解释了1750年持续至今的财富和经济复杂性大爆发，还将商业设计、物理技术、社会技术这三种进化机制纳入到复杂经济学体系中，甚至以复杂性重新定义了经济、商业和财富及其逻辑，堪称一本新的《国富论》。这是一本

全面反思经济学的雄心之作，也是重新认识人性、组织和社会的良心之作。

本力

中国经济学教育科研网主编

过去几十年间，经济思想发生了根本性的变化。埃里克·拜因霍克在这本书中带领我们踏上了一段饱览这些思想的闪耀旅程，非常值得一读。

布莱恩·阿瑟

圣塔菲研究所元老，复杂性科学奠基人

《财富的起源》是一本杰作。拜因霍克通过一系列广泛的研究证明了经济和金融将为何以及如何经历一场思想革命。这本书意义重大，任何对商业和社会产生的影响感兴趣的人都应该读一读。

迈克尔·莫布森

畅销书《反直觉思考》《魔鬼投资学》作者

百余年来最深刻的变化

在我写这本书的时候，经济学正经历着百余年来最深刻的变化。我相信，这种变化代表着世界思潮的重大转变，这将对我们以及后面几代人的生活产生实质性的影响。此外，我还相信，正如生物学在20世纪成为一门真正的科学那样，经济学也将在21世纪成为一门科学。

伟大的历史学家和哲学家托马斯·库恩（Thomas Kuhn）曾指出，科学探索不是在平稳、平坦的进程中展开的。研究库恩作品的一位学者说过，科学的进步中只有“些许平和的插曲，这些插曲是被思想上的激烈革命不时打断的……在每一次革命中，旧的观点将被新的观点取代”。[1]库恩将这些剧变的阶段称为“范式转移”，尽管这个词语因为被滥用而失去了威力，但本书将阐明，我们所见证的今日的经济学，事实上只是这样一次范式转移的初始阶段。

为何经济学的范式转移如此重要？为何我们即便不是经济学家也要关注？大多数人将经济学视为一个枯燥、高深的领域——托马斯·卡莱尔（Thomas Carlyle）给它取了一个响亮的名字：惨淡的科学。[2]经济学的理念之所以重要，是因为它们深刻地注入了社会结构之中。从进行何种类型的抵押，到人们为退休所做的投资，再到我们为谁投票，经济学理念会影响每个人的选择。经济学理念还为商界和政府领导的决策提供了关键框架，这也会对所有人造成影响。就像抽象的科学理论会通过我们乘坐的飞机、服用的药品和使用的电脑而化虚为实，经济学理念也会通过机构、商品和服务以及公共政策而在生活中变成实际的存在。约翰·梅纳德·凯恩斯（John Maynard Keynes）在其著作《就业、利息和货币通论》（*General Theory of Employment, Interest and Money*）的结论部分写到：“经济学家和哲学家的思想无论对错，都比世俗理解的更有影响力。事实上，世界就是由这些思想统治的。”[3]综观历史，糟糕的经济学理念曾导致数百万人遭遇苦难，而优秀的经济学理念则为繁荣奠定了基础。

尽管经济学思想十分重要，但很少有人在学术殿堂外能够意识到今天该领域正在经历根本性的变化。本书是关于我眼中的复杂经济学革命的故事，书中讲述了复杂经济学究竟为何物，揭露了经济学者的一些深

层的奥秘，以及经济学对于整个商业和社会的意义何在。

某些科学革命天生就是完整的，是个别天才的成果，例如爱因斯坦的相对论。另外一些则是许多人经过几十年的努力得出的成果，比如1900年至1930年的量子物理学。复杂经济学革命与后者类似，它是全世界许多研究者多年工作的结晶。尽管复杂经济学里的许多思想有着深厚的历史根源，但直到20世纪70年代晚期，这场革命才开始酝酿，当时物理科学的进步使得一小部分经济学家、社会科学家开始思考是否存在一种全新的方式来看待经济。20世纪80年代至90年代，大规模、成本低廉的计算工具的出现允许研究者利用新式的、不可预见的方式来探索他们的理念，从而推动了复杂经济学革命的发展。

这场革命尚未成功，本身也充满了争议。部分经济学家将会同意本书提出的观点，部分将会坚决地反对，而大部分人将会同意某些观点，同时反对另一些观点。正如一个老笑话所说的那样，如果你想要得到四个观点，只要询问两位经济学家就够了。尽管如此，我仍然相信，思潮正坚定地朝着复杂经济学的方向涌去，并且复杂经济学的概念将在未来数十年为经济学理论和实践奠定基础。本书的目的在于讲述复杂经济学的故事，使之更容易为普通听众接受。与此同时，我还将提出自己的观点，说明这些思想的意义为何，以及如何将它们应用到商业、金融和公共政策的现实世界里。

在写这本书时，我在脑海里设定的是三类读者。第一类是商业领袖、发明家以及政策制定者之中热衷于了解经济和科学将如何影响他们的工作的人。对于这类读者，我想说的是，这不是一本指导你周一上午该做什么的书。复杂经济学包含一套非常有力的概念，它们与高管和政策制定者所面临的大部分棘手的问题是高度相关的。然而，读者将不会在本书中看到龙头企业应用复杂经济学的具体案例，也不会看到什么十步轻松成功法。许多管理学著作承诺读者，它们的理念是最前沿的，同时也被许多公司多年来的成功实践证明是行之有效的——毫无疑问，两者是不可兼得的。我的主张是，复杂经济学确实是前沿的，但这也意味着它没有获得广泛的应用。一些公司已经利用它进行了实验，另外一些公司基于直觉和巧合，无意中也进行了与复杂经济学思想相一致的实践。尽管人们有过一些将这些经验转化为管理工具的尝试，但我认为阐述如何应用这些思想还为时过早。[4]我的关注点不在于案例和工具，而在于帮助读者更新他们关于经济系统如何发挥作用的认知。我希望这本书是一本“周日上午指南”——其目标不在于告诉你具体应该怎么做，而在

于改变你的思考方式。

第二类读者是对经济学、社会组织、公共政策以及科学感兴趣的任何人。见证科学革命的发生，尤其是这样一个将对全世界产生重大影响的科学革命的机会并不易得。对于这类读者，我希望这本书能为他们提供一个前沿视角，来见证这场21世纪最激动人心的变革。本书并没有预设读者拥有经济学或科学背景，相关讨论也不会过于技术化。然而，我会介绍大量对于读者来说并不熟悉的术语。事实上，复杂经济学的其中一个重要贡献就是提供了一套讨论和理解经济学、商业和社会组织的新语言。

第三类读者是学者和学生。对于这个群体，我希望本书能够成为一份有用的非技术性综述，即我们身处何处、将要去向何方。同时，我还希望本书能够澄清（尽管我们肯定无法解决）当前某些关于经济学的争辩，并提出一些新的论题。如果学者和其他读者想要进一步了解相关内容，注释应被当作讨论不可分割的一部分。它们提供了参考，以及在正文中并未呈现的更多技术讨论和延伸。如果读者对于到原始素材中寻找答案感兴趣，我们也将提供完整的参考文献目录。



扫码下载“湛庐阅读”APP，
搜索“财富的起源”，
获取本书参考文献。

目 录

序言　百余年来最深刻的变化
第一部分　范式转移：从传统经济学到复杂经济学
01　财富起源之谜由来已久
02　传统经济学的时代已经结束
03　一场动摇传统经济学根基的辩论
第二部分　洞悉财富起源真相的工具：复杂时代需要复杂经济学
04　定义复杂经济学：审视经济的五种视角
05　视角一：经济是一个动态系统
06　视角二：行为主体是通过归纳推理进行决策的
07　视角三：经济是一个复杂互动网络
08　视角四：经济是一种涌现现象
09　视角五：经济是一个进化系统
第三部分　财富起源于进化：复杂经济学中的三种进化力量
10　力量一：商业设计
11　力量二：物理技术
12　力量三：社会技术
13　经济是三种力量共同进化的结果
14　重新定义财富
第四部分　财富的积累及增长：如何利用复杂经济学理论决胜于商业社会
15　战略：把机构和社会设计成更好的进化体
16　组织：创造“心智社会”，释放个体的创造力
17　金融：预期的生态系统
18　公共政策：左右派争论的终结
后记
注释
译者后记

第一部分

范式转移：从传统经济学到复杂经济学

I

THE ORIGIN OF WEALTH

THE ORIGIN OF WEALTH

宇宙的历史有可能是由少数几个比喻构成的。

——豪尔赫·路易斯·博尔赫斯《迷宫》

01

财富起源之谜由来已久

肯尼亚西南地区的一处偏远乡村中坐落着一座茅草屋，茅草屋最中间的房间烟雾缭绕，我坐在屋里的小窗台上，背靠着粪土砌成的墙壁。茅草屋主人是一位年长的马赛族人，他的脸庞显得睿智、饱经风霜，双目炯炯有神，一直在礼貌地询问关于我的家乡和家庭的问题。

随后，他开始打听我的经济实力。“你有多少头牛？”他的目光透过煮饭的火苗聚焦在我身上。我顿了一下，礼貌地回答：“一头都没有。”我的向导是当地的马赛族教师，已经成了我的朋友，他替我翻译了一下。房间里随之响起了一阵喃喃低语，显然对我这位陌生人充满好奇的村民们正在消化我刚才的回答。思考片刻之后，长者说：“我为你感到非常遗憾。”但他的声音和脸庞除了透露出这种遗憾外，还夹杂着一丝困惑：如此穷困潦倒之人怎么有钱大老远地跑到这儿来，而且还拥有一台相机？

当话题又回到我的家庭时，我提到自己有一位叔叔曾经在马里兰州的农场养了一大群牛。于是，之前的困惑得以解开，他们立刻点头表示理解——很明显，这位客人是仰仗富人叔叔的游手好闲之辈，单纯依靠叔叔的财富（牛群）来旅行和生活。

财富的未解之谜

财富为何物？对于马赛族人来说，财富是以牛的数量来衡量的。对于本书的读者而言，则以美元、英镑、欧元、日元或其他货币作为衡量标准。200多年前，伟大的经济学家亚当·斯密指出，在整个历史进程中，人们用以衡量财富的标准是多种多样的：“在早期社会，牲口被当作商业上的通用媒介，虽然这是最不方便使用的一种媒介……据说食盐是埃塞俄比亚商业和交易的通用媒介；在印度的某些沿海地区，人们使用的通用媒介是某种贝壳；纽芬兰是干制鳕鱼；弗吉尼亚是烟草；美国的部分西印度殖民地是糖；其他某些国家则是牛皮或熟皮；我还听说，在苏格兰的一个村庄，至今工匠去面包房或酒馆的时候带的还是钉子，而

不是钱。”[1]

财富是一种固有的、有形的东西吗？牛、鳕鱼和钉子是因为内含着某样东西而被赋予了价值吗？对于马赛族人来说，财富存在于牛身上，这是人人可见的。牛为他们和他们的家人供应着牛奶、牛肉、牛骨、牛皮和牛角。然而，正如亚当·斯密在其著作《国富论》中所述，财富并不是一个确定不变的概念，事物的价值取决于某个特定时间点上他人的支付意愿。对于马赛族人来说，甚至一头牛的价值在今天和明天都会存在差别。对于用货币衡量财富的人而言，财富更是一个瞬息的概念。发达国家的大多数民众从未见过或触摸过他们的大部分财富——他们辛辛苦苦存下来的钱只存在于银行电脑的电子信号里。然而，这些幽灵般的信号可以通过刷信用卡或点击鼠标的动作，转换成可以触摸的物品，比如牛、鳕鱼、钉子或任何他们希望得到或者有能力购买的物品。

财富最初的起源在何处？人们流下的汗水和脑袋里的知识是如何创造出财富的？为何随着时间的流逝，世界变得越来越富裕？人们是如何从以牛交易转变成以微芯片交易的？这一系列问题最终将把我们导向一个关于财富的最重要的未解之谜：人们如何才能创造出更多财富？我们可以基于狭隘的自身利益提出这个问题，但也可以提出一个更宏大的问题，即如何才能增加社会财富？管理者应该如何发展企业，才能够为人们提供更多的工作岗位和机会？政府应该如何发展经济，以解决贫困和不平等问题？全世界的不同社会应该如何创造资源，解决教育、医疗和其他急迫的问题？全球经济应该如何以可持续的方式发展？财富或许买不来幸福，但贫困却可以“买”到世界上数以百万计的民众的痛苦。[2]

本书旨在探讨的问题包括：财富是什么？财富是怎样产生的？如何为个人、企业和社会创造更多的财富？这些问题都是经济学领域最重要，也最古老的问题。然而，综观历史，经济学似乎很难回答这些问题。本书的观点在于，关于这些基本问题的新答案是在最近几十年的工作中产生的。这些新答案不仅源自经济学家的工作，同样还源于生物学家、物理学家、进化论理论家、计算机科学家、人类学家、心理学家和认知科学家的工作。我们将会看到，现代科学，尤其是进化理论和复杂适应系统的理论，为我们提供了一种全新的视角，来看待这些存在已久的经济问题。

在第1章，我将大致列出本书的主题，简要阐述将在后续章节探究的观点。但在找到一种全新的视角之前，我们必须调整一下自己看问题的角度。对于大部分人来说，经济只是一种日常生活中理所应当的事

物，人们很少去思考它。我们真正思考经济的时候，通常是在一种被普林斯顿大学经济学家保罗·克鲁格曼（Paul Krugman）称为“上涨下降经济学”的情境下，即“股市在上涨”，而“失业率在下降”。[3]事实上，我们需要从经济短期的上涨下降的摇摆曲线中跳脱出来，将经济作为一个整体、一个系统来思考。

人类最复杂的创造

你可以环顾一下自己的房子，看看自己的穿着，望一眼窗外。无论你身处何处，无论是最大的工业城市还是最小的乡村，你都被经济活动及其产生的结果包围着。地球上每时每刻都充斥着设计、组织、制造、服务、运输、沟通、购买和销售等活动。[4]

以上这些活动的复杂性令人难以想象。假设有一座乡村小镇，那是一个安静而简单的地方，是你逃避现代喧嚣生活的去处。镇上的居民将你拥戴为他们仁慈的统治者，赋予你无上权力，但有一个条件——你要确保镇上的居民每天饱有其食，蔽有其衣，居有其所。每个人只会按照你的指令行事（没有指令则一动不动），因此，每天早晨，你必须制作一个工作清单，安排镇上所有的经济活动。[5]你要写下所有必须完成的工作和所有需要协调的事务，以及所有事务的完成时间和顺序。无论是确保威瑟斯本夫人花店的玫瑰到货，还是让纳特利先生的腰痛保单获得赔付，总而言之，这个清单再详细也不为过。即便是一个小镇，这都会是一个太长、太复杂而不可能完成的清单。但这相对于全球的经济管理来说又是一个非常小的体量。全世界每天、每分钟要作出数万亿的复杂协调决策，是它们在维持着世界经济的运转。然而，并没有人负责列出全球的工作清单，也不存在一位仁慈的统治者，由他确保莫桑比克渔民捕获的鱼儿要被运到韩国的餐馆，被当作计算机工作者的午餐，而这位计算机工作者负责制造个人电脑的零件，电脑会被米兰的时装设计师用来设计出被芝加哥利率期货交易商购买的西装。可令人感到惊奇的是，诸如此类的事情每天都在自下而上、自行组织地发生着。

关于经济学第一令人震惊的经验事实莫过于经济确实存在，第二令人震惊的经验事实则是经济日复一日地在运转。全球经济系统为世界上70多亿人口中的大部分人（可惜不是全部）提供了工作、食物、房子、衣服以及从凯蒂猫（Hello Kitty）手提包到医疗激光等的各种各样的产品。如果有人能够联想到其他高度复杂的人造系统，如国际空间站、互

联网，我们可以明显看出，全球经济的复杂程度比其他任何人类创造出来的物理或社会结构的复杂程度还要高出几个数量级。[6]

经济是复杂性方面的一个奇迹，因为没有人设计它，也没有人运营它。当然，首席执行官、政府官员、国际组织、投资人和其他人确实在努力管理经济的某些特定部分，但当人们抽身而出时就会发现，没有人能够掌控它。[7]

经济是如何发展成为今天的模样的呢？科学告诉我们，历史起源于原始状态，即真正意义上“赤身裸体”的状态。人类的直系祖先是早期原始人，他们有着硕大的脑袋和灵巧的双手，他们生活的非洲大草原距离我同马塞族人坐在一起的地方并不遥远。那么，人类是如何从自然状态发展出了令人惊讶的、自组织的、复杂的现代全球经济的呢？

250万年经济史概要

很多人基于直觉认为，人类朝着经济复杂性的迈步是一个缓慢而平稳的过程，是一个从石器到DVD播放机的线性发展过程。而由考古学家、人类学家、历史学家和经济学家共同拼凑出来的实际故事却完全不是这样的，它更富有戏剧性。[8]

故事开始于大约700万年前，那时第一批原始人出现在了地球上，他们的后代非洲南方古猿在大约400万年前学会了直立行走。[9]约250万年前，东非猿人拥有了相对较大的大脑并开始制作粗糙的石器。我们可以认定这些石器是最早的产品，并且可以想象一下，在某个时期，两位原始人祖先坐在非洲大草原的尘土里交易石器。现在人们将大约250万年前的这个节点看作人类经济的开端。随后，直立人又用了将近100万年的时间发明了火，并且开始用石头、木头和骨头制作各种各样的工具。生物学意义上的现代人——大约在13万年前出现的智人，制作出了更加复杂和多样的工具。在某些情况下，智人在何时获得了极为关键的语言技能还存在很大争议。第一批人类的经济活动主要局限于近亲群族中的觅食和基础工具的制作上。

直到3.5万年前才开始出现更加稳定的生活方式的种种迹象，人们有了墓地、洞穴壁画以及装饰品。在这个时期，考古学家开始发现早期人类进行跨群体贸易的证据，包括在墓葬遗址里发现的非本地材料制成的工具，在非海岸部落里发现的贝壳珠宝，以及让人联想到贸易路线的活动方式。[10]贸易的最大好处在于它为劳动分工创造了条件，有记录表

明，在这个时期，工具和人工器物的种类大大增加。正如来自法国图卢兹大学的保罗·西布赖特（Paul Seabright）所言，非亲属间的合作性贸易是人类独有的活动。[11]其他物种都未曾发展出人类经济所独有的、陌生人之间的贸易与劳动分工的组合。实际上，密歇根州立大学的理查德·霍兰（Richard Horan）及其同事认为，智人独有的贸易能力赋予了他们关键的优势，使得他们在与其他古人类物种（如尼安德特人）竞争时得以存活下来，而其他古人类物种却走向了灭绝。[12]

在拥有了固定的居住地点、多种多样的工具以及贸易网络之后，我们的祖先在文化和经济上实现了一定程度的跃升，即人类学家所谓的狩猎采集生活方式。从考古记录中，我们可以或多或少地了解狩猎采集时代的祖先如何生活以及他们的经济形态是什么样的。此外，对于这种生活方式，我们还有另外一个丰富的信息源。现在地球上仍有少数几个与世隔绝的地方存在着狩猎采集部落，这些部落与现代社会的接触寥寥无几，千百年来几乎维持着原样。人类学家则将这些部落看作早期人类生活的实况转播舱。

两个部落的故事

让我们来看看这样两个部落。第一个部落是亚诺玛米族人，他们是一个制造石器的狩猎采集部落，生活在巴西和委内瑞拉偏远边界的奥里诺科河沿岸。[13]第二个部落是纽约人，他们打着电话，喝着拿铁，生活在纽约和新泽西边界的哈德孙河沿岸。这两个部落人群的体内都有着人类共有的大约30 000个基因，因此，从生物学和天生智力的角度来说，他们毫无差别。然而，纽约人的生活方式与亚诺玛米族人保存完好的狩猎采集生活方式可谓天差地别，后者尚未发明车轮，没有文字，他们的计数系统还局限在“1、2和许多”的水平。

如果我们仔细观察两个部落的经济就会发现，亚诺玛米族人的工作机会集中在从森林里采集食物、猎取小型猎物、种植少量的水果和蔬菜以及维护住所上。亚诺玛米族人还会制造一些产品，例如篮子、吊床、石器和武器。他们以村为居住单位，每村40～50人，每个村的村民与该地区的约250名其他村的村民之间会发生货物和服务交易。亚诺玛米族人每年的人均收入约为90美元（这是一个估值，因为他们不用钱，也没有数据统计），而纽约人在2001年的人均收入约为36 000美元，是亚诺玛米族人的378倍。[14]这里对于他们的幸福程度、道德优越程度或与环境

的和谐程度不做评判，但在物质财富上，我们可以看出两个部落之间存在明显的巨大鸿沟。亚诺玛米族人的人均寿命低于纽约人，在亚诺玛米族人的一生中，他们必定会遭遇纽约最贫困人群都不会面对的无常、疾病、暴力、来自环境的威胁以及其他困难。在特定的年份，生活在亚诺玛米村庄的人的死亡率是生活在纽约市的人的8倍。[15]

但是纽约人富有不仅仅是因为绝对收入水平，还因为财富可以让他们买到令人难以置信的、丰富多样的物品。这里你可以设想一下假如自己拥有跟纽约人一样的收入，但只能买到亚诺玛米族人生活中的物品会是什么情形。[16]如果你花36 000美元修缮你的小泥屋，购买村里最好的陶壶，享用最高级的亚诺玛米族饭菜，以亚诺玛米族人的标准来看，你极度富有，但你仍会觉得自己比普通的纽约人贫困，因为他们拥有耐克运动鞋、电视，能到佛罗里达度假。对比之下，普通纽约人的经济选择简直多得惊人。[17]肯尼迪机场附近的沃尔玛超市存货数量超过10万种，有线电视里的频道超过200个，巴诺书店的书超过800万种，当地超市里有275种不同的早餐麦片，普通的百货商店里有150种唇膏，而仅在纽约市就有50 000多家餐厅。

零售商用“库存单位”（SKUs）作为衡量标准，来计算店铺出售的商品种类的数量。比如，5种不同类型的牛仔裤为5个库存单位。如果有人清点亚诺玛米族人所有商品和服务的种类数量，如不同石斧的型号数量、食物的类型数量等，他就会发现，亚诺玛米族经济中的库存单位总数不过数百个，至多数千个。[18]我们无法确切得知纽约人经济中的库存单位数量，但通过一系列的数据源，我粗略地估算出总数约为10^{10}（即数百亿个）。[19]地球物种的总数量约为10^{6}～10^{8}个。因此，纽约人和亚诺玛米族人之间的显著差距，不是用美元衡量的“财富”上的400倍差距，而是纽约人经济与亚诺玛米族人经济的复杂性与多样性之间数百万倍，或者说8个数量级的差距。

亚诺玛米族人的生活方式与1.5万年前我们祖先的典型生活方式相同。[20]这听起来距离我们很遥远，但就人类物种的经济历史而言，亚诺玛米族的世界距离我们非常近。如果我们将第一批工具的出现看作起点，从第一批工具发展到亚诺玛米族式的狩猎采集的经济水平和社会成熟度，用了248.5万年，这在经济史上占了99.4%的时间（见图1-1）。而从亚诺玛米族人的人均90美元、10^{2}个库存单位的经济水平发展到纽约人的人均3.6万美元、10^{10}个库存单位的经济水平只用了人类经济史上0.6%的时间。

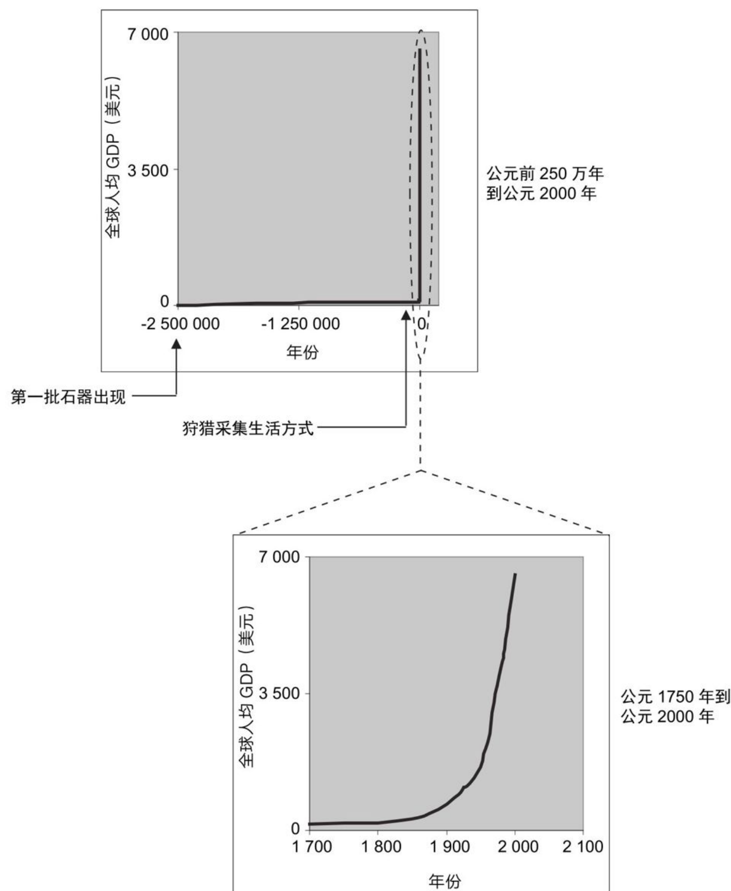


图1-1　人类财富的爆炸性增长

资料来源：布拉德福德·德隆对于公元前100万年到公元2000年期间经济增长的概算，以及公元前250万年到公元前100万年期间经济增长的推测。人均GDP以1990年的国际美元作为计算单位。

假如我们进一步放大并且更加细致地审视过去1.5万年，就会发现一些更加令人感到惊讶的结果。狩猎采集者的世界通往现代世界的经济旅途在1.5万年间的大部分时间里的进程都非常缓慢，但在最后250年出现了爆发性的增长。加州大学伯克利分校经济学家布拉德福德·德隆（Bradford　DeLong）收集的数据显示，从人均90美元的狩猎采集经济社会到公元1000年古希腊时期人均150美元的经济社会，用了1.2万年。[21]

直到公元1750年，世界范围内的人均国内生产总值（GDP）才达到180美元，是1.5万年前我们狩猎采集者祖先的两倍。到了18世纪中叶，在短到不可思议的250年间，世界的人均GDP增长了37倍，达到6 600美元的水平，而最富有的社会，例如纽约，他们的人均GDP还要远远高于

这个数值。[22]全球财富的增长几乎攀升为一条竖直线，时至今日我们仍在沿着这条竖直线往上走。

人类财富的爆炸性增长

我们简要总结一下250万年间的经济发展史：在非常长的一段时间内，几乎什么都没有发生；然后突然之间，天翻地覆。人类历经经济史上99.4%的时间，达到亚诺玛米族人的经济水平；然后又花了0.59%的时间，达到1750年的经济水平；最后经过0.01%的时间，全球财富跃升到现代世界的水平。从另一个角度来说，超过97%的人类财富是在最近0.01%的历史阶段被创造出来的。[23]正如经济历史学家大卫·兰德斯（David Landes）所描述的那样，“比起自己曾孙辈的生活水平，1750年的英国人的生活水平与恺撒时期士兵的生活水平更为接近”。[24]

至此，在对我们面对的现象有了更深入的了解后，我们可以在提问列表中增加更多的问题。

- 像经济这样复杂、高度结构化的事物是如何被创造出来，又如何以自组织和自下而上的方式运作的呢？
- 经济的复杂性和多样性为何会随着时间的流逝而增强？为何经济的复杂性与财富值之间存在关联？
- 为何财富的增加和经济的复杂性会在突然之间爆发，而不是平稳增长呢？

任何试图对财富的定义及其起源作出解释的理论都必须回答这些问题。尽管我们已经从历史叙述的角度知道了经济史上发生的故事，例如农业革命、工业革命等，但我们仍然需要一个理论来阐释经济是如何发生的，以及为何会发生。我们需要一个理论，带领我们探索从生活在原始社会的早期人类到亚诺玛米族的狩猎采集生活方式，再到纽约以及此后的状态。

经济的进化

现代科学正好为我们提供了这样一种理论。本书将论证财富的起源是一个简单有力的三步走公式（即变异、选择、放大）的结果，这也是进化的公式。该程序曾驱动了生物圈的进化次序和复杂性，也驱动了“经济圈”的进化次序和复杂性。[25]并且，同样的程序不仅促成了寒武纪时期物种多样性的爆发，还促成了工业革命时期库存单位多样性的爆发。

我们习惯于从生物学的角度思考进化，但现代进化论将进化视为一个更加广泛的概念。进化是一种算法，是创新的通用公式——通过独具一格的试错法，它可以创造出新的设计，解决难题。进化不仅可以在DNA的“基质”上，还可以在拥有正确信息处理和信息存储特征的系统上发挥作用。[26]简而言之，进化的“变异、选择、放大”的简单处方就像一种计算机程序——一种用来创造新奇事物、知识和增长的程序。进化是某种形式上的信息处理过程，它能在计算机软件、大脑、人类文化和经济等方面开展创建秩序的工作。

经济学与进化论有着一段很长的共生历史。针对该历史，有一种批评意见认为，经济与进化系统相似的类推过于宽泛。例如，有人说，计算机行业就像一个生态位，其中生存着不同的“物种”，包括芯片设计师、硬件驱动制造商、硬盘驱动器生产商、软件提供商等，他们共同参与生态位里的“适者生存”竞争。保罗·克鲁格曼将此类经济和生态系统的比喻式对比称为“生物圣经”。[27]本书提及的大部分研究者都同意克鲁格曼所说的，这种“生物圣经”既不是好科学，也没有太多的启发意义。如今研究者在将经济理解为一个进化系统时，会自觉避免类似的比喻，转而将注意力集中在理解进化的通用算法是如何植入人类经济活动的信息处理基质上的。由于生物和经济系统共享了进化的核心算法，所以它们之间存在某些相似性，但它们实现进化的方式其实存在很大的差别，必须在各自的情境下理解。

从科学的角度来说，将全球经济看作一个进化系统的比喻与按照字面对此进行理解，这两者之间的差别非常关键。经济系统与生物系统相似的说法并没有为我们提供什么具有科学价值的东西。但是，在一个更加普遍而广泛的进化系统类别里，经济系统和生物系统都是它的子类，这一说法给我们提供了很多东西，因为研究者相信进化系统里存在着一些普遍规律。[28]科学家认为，自然的某些特征是普遍性的。比如，重力在地球上和在宇宙中的作用是一样的，它对于原子、苹果和星系的作用也毫无差别。现代进化理论家认为，同重力一样，进化也是一种普遍现象，也就是说，无论是在生物DNA的基质，还是在电脑程序、经济系统

及遥远星球外星生物的基质里运行进化的算法，进化都遵循着某些特定的普遍法则。

如果经济真的是一个进化系统，并且进化系统存在某些普遍法则，那么我们可以得出一个结论：经济上也存在某些普遍法则——这就与许多观点发生了冲突。这些观点认为，经济领域并不存在普遍法则，我们永远不可能对经济作出精准预测，但是在未来的某天，我们可以获得对经济现象的更深刻的理解。这也意味着，在未来，经济学或许可以为商业和公共政策给出精准的建议，并具备前所未有的科学权威。

有人或许会说，更加科学的经济学前景将会激动人心，并且能为世界带来许多潜在的利益。其他人或许会认为这是将科学应用到人类社会问题上的又一次被误导的尝试。这些批评让我们回想起19世纪晚期和20世纪初期，社会达尔文主义运动中产生的一些令人反感的观点。当时的一些哲学家，如赫伯特·斯宾塞（Herbert Spencer），试图通过粗暴的比喻，将达尔文理论应用到社会和经济领域。[29]社会达尔文主义者用“适者生存”法则（人们经常误认为这是达尔文提出的，其实这句口号出自斯宾塞）来为阶级不平等、种族歧视、殖民主义和其他社会不公平现象作辩护。我们即将讨论的经济进化的新观点与社会达尔文主义的旧观点没有任何共同之处。事实上，它们的意义正好相反，合作与“适者生存”的个人主义在经济发展中同样重要。关于社会工程的告诫已经受到了重视，后续讨论的新理论将能够帮助我们揭示为何经济现象难以预测，以及为何历史上的大型社会工程都会以失败告终。

创造出适合的设计

什么样的算法才是进化？它能起到什么作用？哲学家丹尼尔·丹尼特（Daniel Dennett）[(1)]将进化定义为一种用来创造“不用设计师的设计”的通用算法。[30]举个例子来说，常见的陆生蚯蚓就是为了能在北美和欧洲的森林、草地和家庭花园环境中生存和繁殖而存在的巧妙设计。它本质上是一种管状生物，这种结构能让它在泥土中穿梭，从一端摄入泥土并从另一端排出，在此过程中吞食大量富有营养的微生物、获取足够的卡路里，以便能够找到更多食物，繁殖后代。这种特殊的设计布满了接触传感器和震动传感器，可以帮助生物体远离捕食者。生物身体的大部分身段上都有备用系统，如果被切成两段，它还可以自行再生。陆生蚯蚓还可以大量繁殖，以便提高后代存活和繁殖的概率。这种聪明设计是由

进化算法创造出来的，无须理性设计师的存在。[31]

进化创造了设计，更确切地说，借由试错的过程，进化发现了设计。许多候选设计被创造出来，然后被环境淘汰。成功的设计得以保留、复制和存续，失败的设计则会被摈弃。经过反复试验，适合特定目的和环境的设计就会被创造出来。如果条件正确，设计方案之间为获取有限资源的竞争会使组织结构日益增大，复杂性日益增强，这是因为进化是建立在过去为未来创造出更加新颖的设计的基础之上的。[32]随着世界的变迁，进化创造出来的设计也通常会朝着更加聪明、甚至更加惊人的方向变化。进化是一种在大量、几近于无穷的可能设计中寻找几近于无穷小的"适宜"特定目标和环境的设计方法。就好像丹尼特所说的那样，进化是一种搜索算法，像是"在草垛（可能性）中搜寻细针（好的设计）"。[33]

有些人或许会用"没有设计师的设计"来解释生物进化，但为何在拥有许多人类设计师的情况下，我们需要用"没有设计师的设计"来解释经济领域中的财富创造呢？难道我们不是创造人类经济的上帝吗？我们习惯于将人类的理性和创造力看作创造财富的主要驱动力，毕竟财富是由聪明、善于创新的人创造出来的，他们为产品和服务想出新的创意，又为生产和销售做了许多努力。我将在后续章节中论证，人类的理性和创造力的确在财富创造的过程中发挥了重要作用，但事实与我们通常所认为的有些偏差。理性和创造力助长塑造了进化算法对于经济所起的作用，但并没有取代它。

想想看你穿的衬衫、裤子或其他衣物——它们的设计由何而来？[34]你或许会说，很明显啊，是服装设计师设计的。但事实并非仅仅如此，事实或多或少地更接近于以下这种情况：许多设计师在衬衫应该是什么样的已有概念上，通过发挥他们的理性与创造力来创造各式各样的"衬衫"，并将它们的样子勾画出来。随后，这些服装设计师会查看设计图，选出他们认为顾客会喜欢的一些设计，并制作少量的样品。接下来，设计师把这些样品展示给服装公司，服装公司又选出一部分他们认为顾客会喜欢的设计，并安排生产。服装公司会向不同的零售商展示这些产品，零售商同样会选择一部分他们认为顾客会喜欢的设计。有了订单之后，服装公司开始批量生产，并向零售商供应这些衬衫。最后，你走进商店，浏览各式各样的衬衫，选择你所喜欢的并付款。不同的设计、根据某种标准而作出的选择以及对于成功设计的量产，这一切不仅发生在服装公司里，也发生在整个服装市场。你的衬衫并不是设计出来的，它

是进化而来的。

为何时装行业要经历这些迭代，甚至可以说浪费的过程呢？你的衬衫是进化出来的而非设计出来的原因在于，没有人可以从接近于无穷多的衬衫设计中，准确预测你所喜欢的衬衫的样子。我们将会看到，尽管人类理性中蕴含着能力和美德，但在短时间内对经济这样复杂的系统进行预测几乎是不可能成功的。在作出经济决策时，我们会尽最大努力发挥大脑的优势，然后进行试验，不断探索向着未来前进，保持和维护有用的部分，同时摒弃无用的部分。我们的目标、理性和创造力确实在经济中起到了作为驱动力的重要作用，但它们只是作为更广阔的进化过程中的一部分而起作用的。

经济进化并不是一个单独的过程，而是三个相互联系的过程。第一个过程是技术的进化，这是历史上经济发展的关键因素。1750年左右出现的经济急速增长与工业革命的技术飞跃恰好吻合。但技术的进化只是一部分。来自哥伦比亚大学的进化经济学家理查德·纳尔逊（Richard Nelson）曾指出，事实上有两类技术在经济增长的过程中扮演了重要的角色。[35]第一类是物理技术，也是我们通常所说的技术，例如制铜技术、蒸汽机和芯片；第二类则是指人们组织活动的方式，例如农业、法律、货币、股份制公司以及风险投资。纳尔逊认为，物理技术曾对社会产生了重大影响，而事实上社会技术也起到了同样重要的作用，两者是共同进化的。[36]工业革命期间，理查德·阿克赖特（Richard Arkwright）发明了精纺机（物理技术），使得大型工厂里服装生产的方式更加经济（社会技术），这又反过来刺激人们在应用水力、蒸汽和电力进行制造的过程中作出了许多发明。[37]农业、工业和信息革命的历史故事大部分都是物理技术和社会技术相对起舞的历程。

然而，物理技术和社会技术的共同进化也只占了2/3的内容。技术本身只不过是理念和设计。服装精纺这种物理技术本身不是服装精纺技术，事实上是有人需要它成为这种技术。同样，工厂这种社会技术本身并不是工厂，而是有人需要它成为工厂这种组织形式。为了让技术对世界产生影响，有些人或者有些群体需要将物理技术和社会技术从概念转化为现实。在经济领域，这种转变是经由商业实现的。商业将物理技术和社会技术融为一体，并将它们转化为产品和服务带入现实世界。

商业本身就是一种设计。商业的设计包含了商业策略、组织结构、管理流程、文化以及许多其他因素。通过变异、选择和放大的过程，商业设计在时间的进程中不断进化，而市场是对适宜性的最终仲裁者。本

书的其中一个主题，就是论述物理技术、社会技术和商业设计的三向共同进化塑造了我们今天所见到的经济波动和增长模式。

复杂经济学

经济是一个进化系统，这个概念是具有革命性的，尤其是因为它与过去100多年间发展出来的大部分经济学的标准概念是相冲突的。然而，它远不只是一个新颖的概念。进化理论与经济学有着一段很长的紧密相连的历史。[38]事实上，是一位经济学家激发了达尔文某部分最重要的洞见。1798年，英国经济学家托马斯·罗伯特·马尔萨斯（Thomas Robert Malthus）出版了一部名为《人口论》（*As Essay of the Principle of Population*）的作品，在书中，马尔萨斯将经济描述为一场竞争性生存挣扎以及人口增长与人类生产力增长之间的持续斗争。马尔萨斯预言，在这种斗争中，人类将会败北。达尔文阅读了马尔萨斯的著作，并在自传中记录了自己的感想：

> 1838年10月，那是我开始系统查询资料后的第15个月，我在闲暇时无意中读到了马尔萨斯的《人口论》一书，这才充分理解了生存竞争——通过对动物和植物习性的长期观察发现，生存竞争普遍存在。这立即启发了我，即在这样的情况下，有利的变化将被保存下来，反之则被淘汰，这个过程的结果就是新物种的形成。
>
> 由此，我终于找到了一种可行的理论。[39]

由此可见，达尔文对于自然抉择在进化中所起到的关键作用的伟大洞见是被经济学激发的。[40]在达尔文发表进化论后不久，思潮又从进化理论家流向了经济学家。1898年，经济学家索尔斯坦·凡勃伦（Thorstein Veblen）写下了一篇至今看来仍然精彩绝伦的文章。文章指出，经济是一种进化系统。[41]随后不久，现代经济学的奠基者之一阿尔弗雷德·马歇尔（Alfred Marshall），在其知名作品《经济学原理》（*Principles of Economics*）的引言中写到，“经济学家的圣地存在于经济生物学领域”。[42]在随后的几十年里，许多伟大的经济学家，包括约瑟夫·熊彼特（Joseph Schumpeter）和弗里德里希·冯·哈耶克（Friedrich Hayek），都曾致力于探究经济学与进化论之间的关系。[43]1982年，理查德·纳尔逊和悉尼·温特（Sidney Winter）发表了一部具有划时代意义

的作品：《经济变革的进化理论》（*An Evolutionary Theory of Economic Change*）。这是首次将进化论、经济以及最近发展起来的计算机模拟工具联合起来的重大尝试。[44]

尽管经济学界的学者作出了种种努力，但进化论对于主流经济学理论的影响微乎其微。大致从达尔文出版《物种起源》的时间点起，经济学转向了一条殊途。从19世纪晚期开始，经济学的组织框架认为，经济是一个均衡系统，其本质上是一个静止的系统。在后续章节中我们将会看到，从19世纪晚期到20世纪中期，经济学家的主要灵感来源不是生物学，而是物理学，尤其是运动和能量物理学。传统经济学理论将经济看作一个个在大碗底部滚动的橡皮球。最终，橡皮球会在碗底的静止点或均衡点停下。橡皮球将保持不动，直到外部的力量撼动、压弯或冲击大碗，从而将球推到一个新的均衡点。在过去数百年间，经济学的主流框架认为，由于受到技术、政治、消费者品位转换以及其他外部因素的冲击，经济是一个不断地从一个平衡点移动到另一个平衡点的系统。

在经济学家不断追求经济学作为均衡系统这一愿景的同时，在19世纪下半叶，物理学家、化学家以及生物学家对一类远远称不上均衡的、动态而复杂、永不停歇的系统的兴趣日益浓厚。从19世纪70年代开始，科学家们开始将这类系统称为复杂系统。我们将在后续对该术语进行更加细致的讨论，简而言之，复杂系统是指一个由许多动态关联的部分或粒子组成的系统。在此类系统中，部分或粒子在微观层面的互动将导致宏观层面行为方式的涌现。比如，单个水分子处于静止状态时是相当无聊的，但如果将数十亿的水分子放在一起，并适当地赋予一些能量，就会出现复杂的微观漩涡模型。[45]漩涡模型是单个水分子动态互动的结果。如果只有一个水分子，就不能形成漩涡，漩涡是系统本身的集体或“涌现”的产物。

20世纪70年代，随着科学家们对复杂系统行为的了解越来越深入，他们越来越有兴趣地关注到，粒子并不是像水分子那样具有固定行为的简单事物，它们拥有一定程度的智力，能够适应自身所处的环境。水分子无法调整自己的行为，蚂蚁却可以。以人类的标准来看，蚂蚁或许不是非常聪明，但它可以处理从其他蚂蚁那儿和环境中得来的信息，并相应地调整自己的行为。跟单个水分子一样，单个蚂蚁本身不太可能存在。然而，如果你将几千只蚂蚁放在一块儿，它们彼此之间就能互动，通过化学信号进行交流，还能够协调相互之间的活动，例如建造结构复杂的蚁穴，对进攻者组织复杂的防御。科学家将这些具有处理信息和调

整行为能力的部分或粒子称为行为主体，并将行为主体互动的系统称为复杂适应系统（complex adaptive systems）。[46]复杂适应系统的案例还包括身体免疫系统的许多细胞，生态系统里的诸多有机体，以及互联网的大规模用户等。随着20世纪80年代低价、高性能计算机的出现，科学家在理解自然界中的复杂适应系统以及将这类系统看作具有许多共同行为特征的通用类别上取得了较大的进展。实际上，许多生物学家都将进化系统视为复杂适应系统中的一个特殊类别或子类别。

此后，社会科学家越来越怀疑经济是否也是一种复杂适应系统。经济较明显的特征之一就是它是人们相互之间以复杂方式进行互动、处理信息以及调整自身行为的综合体。20世纪80年代至90年代早期，研究者开始对有别于传统模式的经济现象进行实验。[47]他们认为经济不是一个静态均衡系统，这些模式显示，经济是一系列嘈杂的动态活动。就像漩涡源自互动的水分子一样，这些模式表明，繁荣与萧条的复杂模式以及创新的浪潮是经由受刺激的行为主体互动产生的。在过去的几十年里，人们对于将经济理解为复杂适应系统的兴趣和研究急剧增多，在本书中，我们将对这些工作进行回顾。

我将用复杂经济学一词指代相关的主要工作（该词语的创造者是布莱恩·阿瑟［Brian Arthur］，他曾任职于斯坦福大学和圣塔菲研究所）。[48]我们不该根据字面意思认为现在存在一种统一、综合的复杂经济学理论。我使用这个词语，是为了涵盖一系列本书将要探究的理论、假设、工具、技术和现象。在现阶段，复杂经济学是一项正在进行中的工作，或者说是科学哲学家所说的“项目”，而非一个统一的理论。[49]

前方的线路图

如果经济确实是一个复杂适应系统，那么这代表着四层重要的隐含意义。

第一，这意味着在过去的一个世纪里，经济学家对经济的归类从根本上来说是错误的，而且教科书、管理学思想和政府政策中所反映的主流经济学要么是错的，要么最好的情况也只不过是接近正确。

第二，将经济看作复杂适应系统，能够为我们解释经济现象提供一套新的工具、技术和理论。

第三，这意味着财富必然是进化过程的产物。正如生物进化使原始

汤变成了复杂的有机体和生态系统，经济进化使人类从原始状态发展到了现代的全球经济，同时为世界带来了秩序、复杂性和多样性。

第四，历史证明，每当经济学理论出现重大的范式转移，学术界的震动要在很久之后才能传导到外界。在19世纪，亚当·斯密的思想对于自由贸易的发展产生了重要的影响；在20世纪初期及中期，马克思的愿景激发了革命以及社会主义的兴起；在20世纪最后的十几年里，英美新古典经济学的思潮与全球资本主义的上升正好契合。我们还需要数十年才能弄清复杂经济学的社会和政治影响。尽管如此，复杂经济学的轮廓已经充分形成。

我们将在最后获知一个乐观的消息：如果人们能够更好地理解财富的起源，就能利用这些知识找到为人们创造经济增长和机会的新办法。复杂经济学并不是解决一切管理挑战或社会问题的万金油。对自然现象更加科学的理解为创造更好的人类生活条件作出了巨大贡献，同理，对经济现象的科学理解也有可能给全世界人类带来福祉。

02

传统经济学的时代已经结束

时值1984年，约翰·里德（John Reed）遇到了一个难题。[1]在45岁这年，里德刚刚当选为世界顶尖公司花旗银行的董事长和首席执行官。然而，里德所接手的公司刚刚经历了一场重创。在整个20世纪70年代，花旗银行同美国其他主要银行一道，向发展中国家的政府发放了大量贷款，尤其是拉丁美洲的政府。里德的前任沃尔特·里斯顿（Walter Wriston）宣告这些贷款是“安全的银行业务”，因为主权政府从不拖欠债务。里斯顿后来被证明大错特错——1982年8月，墨西哥政府宣告无法偿还它所背负的巨额债务。

一连串类似事件最终导致了全球金融危机的爆发。在随后的几年里，有多个国家出现了大范围的违约、货币贬值和经济崩溃。尘埃落定后，数以百万计的穷困民众发现他们更加贫穷了，银行则发现3 000亿美元从它们的资产负债表上蒸发了。仅花旗银行一家就在一年内损失了10亿美元，另外还产生了130亿美元的坏账。

里德试图弄清究竟发生了什么，是如何发生的，以及如何防止它再次发生。花旗银行和其他所有大型银行里的优秀人才为何会严重地误判了相关的风险？为何没有任何人能够预测到这些贷款将导致的问题？为何在墨西哥发生的局部事件会升级为一场全球危机？为何全世界其他政府的应对措施丝毫不起作用？

里德咨询了各种专家，包括来自学术界、华尔街和各国政府的顶尖经济学家。自就读于麻省理工学院起，里德就对经济学产生了浓厚的兴趣。如果说回答这类问题是某些人的工作职责，那么这些人必定是经济学家。然而，经济学家对于这场危机却给不出任何新颖的或有用的说法。事实上，里德认为他们在危机中给出的建议错得离谱。据科学作家米歇尔·沃尔德罗普（Mitchell Waldrop）所言：“里德判定，在面对世界金融市场时，专业的经济学家毫无用处……里德认为我们或许需要一种全新的经济学方法。”[2]

对于新方法的需求

里德并非唯一一个质疑经济学状况的人。在过去几十年里，人们对于经济学理论的批评声不绝于耳。[3] 1996年，约翰·卡西迪（John Cassidy）在《纽约客》杂志上撰写了一篇极具争议而阅读者众多的文章：《经济学的衰落》。[4]卡西迪指出，经济学已经沦为象牙塔中高度理想化的理论，未经数据检测，充满着不现实的假设。他声称，经济学已发展成一场“大型的学术游戏”，经济学家在游戏里写文章互相吹捧，炫耀各自的数学才华，但毫无兴致证明他们的理论与现实世界的关联。卡西迪认为，大部分企业已经对经济学表示绝望，他还指出IBM、美国通用电气等公司已经关停了它们的经济学部门。

不仅商业人士和记者在批评当前的经济学状况，经济学家本身也是其最严厉的批评者。[5]在《纽约客》的文章中，卡西迪引用了美国总统经济顾问委员会前主席、诺贝尔奖获得者约瑟夫·斯蒂格利茨（Joseph Stiglitz）的说法：“任何看到这些模型的人都会说，它们没有很好地描述现代社会的状况。”[6]在同一篇文章中，也曾任经济顾问委员会主席一职、来自哈佛大学的格雷戈里·曼昆（Gregory Mankiw）同样指出，鉴于经济学家的有用产出较低，应削减他们的研究经费，曼昆还将经济学家比作获得过度补贴的奶农。美联储前主席以及在经济政策制定方面受到高度尊崇的艾伦·格林斯潘（Alan Greenspan）[(2)]曾对他的同事说：“我们真的不知道（经济）是如何发挥作用的……旧模型已经不起作用了。”格林斯潘在更早之前还说过：“许多经济学家无法区分我们所创建的经济模型与现实世界。”[7]

尽管人们对经济学理论的不满愈加严重，但只有很少几位批评者认为这个领域是完全无用或是没有影响力的。相反，大部分商业领导者、政策制定者，甚至长于自我批评的经济学家都承认，经济学产生了一些极其强大、有影响力的理念，这些理念的范围涵盖了从市场效率到自由贸易的好处，再到个人选择的重要性。经济学理论成功与否的衡量标准可以在七国集团中的富裕经济体里被找到，因为在七国集团，这些理念被运用的程度远远超过世界上其他任何国家。现在，七国集团大部分民众认为理所应当的许多理念，例如使用利率调节通货膨胀，货币和财政政策可以抑制经济循环、积极鼓励竞争，社会保障体系可以磨平市场体系的粗糙边缘，产品安全、环境、劳动法规可以保护人们不受市场失灵的伤害——这些理念在100年前、50年前，甚至20年前都不常见，都是学者型经济学家经过一个世纪的艰苦工作发展起来的。

更确切地说，问题在于学术、商业和政策制定者群体越来越觉得经

济学没有发挥它作为一门科学的真正潜质。[8]如今，该领域许多“伟大思想”的历史已经超过100年，太多正式的理论和数学模型要么被不现实的假设拖了后腿，要么与现实世界的数据直接矛盾。问题的关键不在于诋毁过去的贡献，而在于承诺“经济学可以做得更好”，是时候往前走了。

第2章和第3章将为大家揭示为何我们需要一种全新的办法。在一开始，我们将为经济学领域的传统观点或者说我所谓的传统经济学作出定义。我们将快速回顾一下传统经济学理论的历史和重要概念，并在第3章听一听批评者们的整体言论。可以想见，我们只能对200多年的经济学理论走马观花，而忽略掉许多重要的工作。然而，这两章的目的并不在于为传统经济学或全面的批评考察提供教科书般的叙述（更加完整的叙述请参考注解和参考书目），[9]而是在于强调现代经济学发展过程中的一些突出问题，检测它们的优势与不足，并为第二部分中关于复杂经济学的讨论打下一些基础。我们将会看到，为了能使大家透彻地理解经济学在未来的走向，在此之前了解它的过去是非常重要的。

定义传统经济学

我将使用传统经济学一词指代过去100年间占据经济学理论主导地位的一系列理论。至此，我有必要定义这一词汇所蕴含的意义。总的来说，传统经济学是指人们在大学课本里发现的、在媒体上讨论的以及在商业公司和政府大厅里提及的经济学，主要是指学术经济学的主流观点。[10]为了能够更加准确地定义这个概念，我将遵从两大传统经济学主要批评家的引导——哥伦比亚大学的理查德·纳尔逊和宾夕法尼亚大学的悉尼·温特，并使用经济学文献本身作为我下定义的基础。[11]

> 传统经济学是指本科和硕士研究生教科书里阐述的一系列概念和理论。它还包括同行评议调查里声称或假设的、领域内普遍认可的概念或理论。[12]

教科书代表了专业的主流观点，包括进入该领域必须知道的基础观点。[13]但教科书不可避免地会遗漏一些更加高级的资料。因此，我在自己的定义中增加了一些调查文献和文章，它们及时地对某些时间点上经济学领域的状况进行了总结。当然，教科书和调查资料的局限性都在于它们通常将注意力集中在传统的而非前沿的方面，[14]但是这种局限性是有意为之的。我所说的传统经济学是指经济学的历史核心，是那些诺贝

尔奖已经嘉奖的内容。[15]被我归纳在“传统”这一标签下的理念，通常是指经济学家所谓的新古典经济学，在本章的后续内容中，我们将对这个词语作出定义。[16]

判断内容是否能够划归到传统标签之下，必然会有一定的主观性，并且还会存在一些灰色领域——某些理念既可划入又可划出。尽管如此，这个标签仍将在后续的讨论中起到重要作用，因为我们在历史框架和复杂经济学新理念之间划清了界限——这将是一个挑战。在了解这些说明后，既然已经为传统经济学的盒子贴好了标签，那我们就来一窥究竟吧。

大头针制作与看不见的手

我们将从知名度最高的经济学家亚当·斯密谈起。亚当·斯密并不是历史上第一位经济学家，有资料记载，这一荣耀属于古希腊哲学家色诺芬（Xenophon）；“经济学”一词源自他的作品《经济论》（*Oikonomikos*），但亚当·斯密的巨大影响力为我们的讨论提供了一个合适的起点。[17]亚当·斯密出生于苏格兰爱丁堡附近一个名为克尔卡迪的小镇，在世时间为1723—1790年，这属于历史学家所谓的古典经济学理论的阶段（1680—1830年）。[18]亚当·斯密曾就读于牛津大学，但他职业生涯的大部分时间都是在格拉斯哥大学度过的。他的第一部重要著作《道德情操论》不是关于经济学，而是关于道德哲学的。这本书出版于1759年，使得亚当·斯密在相对年轻的年纪就成了苏格兰启蒙运动的一个关键人物。在格拉斯哥大学就读期间，他引起了一位年轻苏格兰公爵的注意，那位公爵带他进行了一趟报酬丰厚的私人旅行。亚当·斯密跟随公爵到了法国，在那里，他接触到了当时欧洲大陆上热议的经济学理念，尤其是重农学派的理念——当时有一群知识分子坚持认为：政府应该限制对经济的干扰，让市场发挥主要作用。从公爵那儿获得的收入给了他经济上的保障，回到克尔卡迪后，亚当·斯密与母亲一起过了6年深居简出的日子，专职写作《国富论》（*Wealth of Nations*）的手稿。该书于1776年出版，并立即被认定为一部伟大的作品。

历史上，经济学家一直对两个基本问题争论不休：一是财富是如何被创造出来的，二是财富是如何分配的。[19]在《国富论》中，亚当·斯密回答了这两个问题。[20]他对于第一个问题的回答简单而有力：人们从环境中获取原材料，经过劳作将这些原材料转变为人类所需的东西，经济

价值由此产生。比如，制陶工人从地里取出黏土并用它来制成一个碗。斯密的伟大洞见在于他提出财富起源的秘密就藏在劳动生产力的提升之中。制陶工人在一个小时内能够制造的碗越多，他就越富有。更高生产力的秘诀则在于劳动分工以及劳动分工所带来的专业化。[21]亚当·斯密引用了广为人知的一个大头针工厂的例子，他观察到，工厂里有10名工人，每人专门负责大头针生产过程的一到两个步骤。[22]这种专业化和合作使得团队每天能够生产48 000枚大头针，或者说每人每天能够生产4 800枚大头针。如果没有劳动分工，工厂里的工人每人每天只能生产20枚大头针，而技艺最差的工人一枚都生产不出来。

人口的增长能够增加社会的总财富，因为可用的劳动力增多了。但人均财富的增加（从而提升个人的生活水平）需要提升生产力，而生产力的提升需要专业化分工。这个逻辑将亚当·斯密引向了第二个经济问题：是什么决定了财富和资源在社会中的分配方式？创造财富需要专业化分工，而专业化分工需要贸易——毕竟，大头针工人不能以大头针为食物，他们必须用它们来交换生活所需的其他物品。但如果大头针工人、农民、渔民、木匠以及其他生产者在同一个经济体中交换他们的产品，什么能够决定物品的分配方式呢？多少个大头针能换一蒲式耳(3)小麦？多少条鱼能换木匠的一把椅子？大头针工人和渔夫谁会更加富有？亚当·斯密作为一名道德哲学家，他自然而然产生的疑问不仅仅是资源是如何分配的，还包括它们应该如何分配，以及对于个人和整个社会来说，什么才是公平或公正的资源分配方式。

亚当·斯密认为，最公正的分配资源的机制是能够让人们追求自身利益、自主决策的机制。毕竟，每个人都是自身幸福的最佳判断者。与此同时，对于社会整体来说，最好的资源分配方式是将资源分配到利用效率最高的地方，由此实现社会总财富的最大化。在道德层面而言，浪费资源是不应该的（对节俭的苏格兰人来说尤其如此），因为这样会减少社会可用的总财富。亚当·斯密提出的基本原理源自他在格拉斯哥大学的导师弗朗西斯·哈奇森（Francis Hutcheson），哈奇森曾呼吁“为多数人争取最大的幸福”。[23]亚当·斯密对于如何实现这个目标提出了一个革命性的观点（至今对于部分人来说仍然如此）：自由竞争是分配社会资源最公正的机制。他指出，如果人们可以自由地贸易，那么谋取自身利益将驱动人们提供他人所需的商品和服务：“我们不能期望从屠夫、酿酒师或银行家的善心里得到晚餐，而要从对他们自身利益的尊重里得到。”[24]此外，利益动机和竞争将一同驱使人们以最高效的方式提供商品和服务：“每个人都会尽己所能，去寻找自身最具优势的工作来赚取他们所能

赚到的利益。”[25]

亚当·斯密指出，对于自身利益的追求将反过来给整个社会带来好处：“（商人）只在意自己的收益，在这种情况下，或是许多其他情形中，他们都被一只看不见的手指挥着，实现了一种与他本意无关的目的……商人追求自身利益的过程时常会伴随社会利益的提升，这比他真正想要提升社会利益更有效。”[26]

这只“看不见的手”为社会带来了资源高效分配的喜人结果，它是竞争市场所产生的机制。亚当·斯密描绘了价格是如何在生产者和消费者于市场相遇的过程中提供这种关键机制的。[27]如果对于有效需求的供应太少，价格就会上涨，生产者就会加大生产，而消费者会减少消费。如果对于有效需求的供应过多，价格就会降低，生产者就会减少生产，而消费者会增加消费。在某个时刻，市场价格会稳定在这两种相反力量的均衡点上：供应等于需求，市场出清。亚当·斯密指出，如果放任不管，人对自身利益的追求和竞争市场将自然而然地将经济引导到这个均衡点上。我们可以引用20世纪80年代的电影《华尔街》里的角色戈登·盖柯（Gordon Gekko）的话，用现代话语来重新阐述亚当·斯密的观点：“贪婪是好事。”

健康的均衡

经济在其自然发展过程中会达到一个均衡点，这种概念可以追溯到亚当·密斯之前，追溯到经济学发展的最初阶段，而且它现在仍是传统经济学的核心概念：对有限资源的竞争必然意味着在经济中存在着相互抗衡的力量或紧张局势。对于17世纪的爱尔兰金融家理查德·坎蒂隆（Richard Cantillon）而言，经济中最主要的竞争局势莫过于人口与土地的粮食产能之间的竞争。坎蒂隆相信，人口过量与饥饿之间的严酷机制将导致经济系统中的工资和物价自发调节，使得人口过量与饥饿最终达到某种平衡。[28]18世纪法国知识分子弗朗索瓦·魁奈（François Quesnay，重农学派的领导人，斯密在法国期间的大部分时间都是同他在一起）则认为，经济领域最主要的矛盾是农业、制造业以及拥有土地的贵族之间的矛盾。在其代表作《经济表》（*Tableau Economique*）中，魁奈声称可以计算出能使经济达到均衡状态的物价和生产水平。[29]对于魁奈而言（他在进入经济学领域之前曾是一名医生），均衡的经济就是健康的经济，就像18世纪的医学观念，如果身体的各种“体液”是均

衡的，那么身体就是健康的。对于亚当·斯密而言，经济领域最主要的矛盾是消费者与生产者之间的矛盾，要达到的均衡是供求之间的均衡。我们应该知道，亚当·斯密所说的供应和需求并不完全等同于如今教科书中的供求概念，如今的供求概念是后来由约翰·斯图尔特·密尔（John Stuart Mill）和阿尔弗雷德·马歇尔创造出来的。

亚当·斯密描述了市场在达到供求均衡的过程中所起的作用，但他并没有详细描述决策的过程，即在此过程中，受自我利益驱动的生产者决定供应多少产品，或者受自我利益驱动的消费者决定购买多少产品。决策过程的核心理念其实来自亚当·斯密的两位同辈：雅克·杜尔哥（Jacques Turgot）和杰里米·边沁（Jeremy Bentham）。

雅克·杜尔哥是路易十五政府的一名部长，是自由放任主义（laissezfame）的知名拥护者——自由放任主义是指政府应该在市场运作的过程中将其干预降至最低的一种思想。[30]尽管杜尔哥持有这样的观点，当时的法国政府还是极大程度上介入了经济的运作（换汤不换药），而杜尔哥作为部长的职责之一则是处理食物短缺问题。[31] 1767年，杜尔哥观察到，如果农民仅仅是将种子洒在地里，就只能收获很少的庄稼。但如果在播种前哪怕只翻一次土，产量就会高得多。如果翻两次土，他的收获不仅会翻倍，甚至能达到原来的3倍。农民在土地上下的功夫越多，他们的收获就越多。但在到达某个顶点后，地力就会衰竭，农民投入的额外单位努力的产出就会减少。在此观察的基础上，杜尔哥提出了收益递减规律（law of diminishing returns）。在大多数的生产过程中，无论是农业、制造业还是服务业，随着某一要素（如劳动力、原材料或机器）投入的不断增多，到达某个时间点后，单位投入所能得到的产出就会越来越少。收益递减规律是使得经济达到均衡的一种关键力量。针对市场的某一价格，生产者会不断增加投入、扩大生产，直到产出不抵投入，也就是说直到生产的增量成本大于增量收入。因此，农民会按照市场的需求量进行生产，不多也不少。如果农作物的价格上涨，农民会更加努力地耕地（或耕种更多土地）；而当价格下降，他们就会减少耕种。如果投入所能得到的产出不会下降，农民会继续扩大生产——无限投入，而这是一种非常荒谬的结果。[32]将生产者成本与供求关系中的供应方面联系起来，杜尔哥的规律为此提供了一种至关重要的理念。[33]

几乎是在同一时间，英国哲学家杰里米·边沁在需求方面也作出了同样重要的贡献。边沁1748年生于伦敦。他是一个神童，4岁开始学习拉

丁语，12岁进入牛津大学。[34]同亚当·斯密一样，边沁将自己定义为道德和政治哲学家。亚当·斯密指出，人类的利己主义是驱动经济发展的动力，但他没有过多提及利己主义是如何转化成具体的经济决策的。边沁认为，对于自我利益的追求是一种基于愉悦和痛苦算计的理性活动。[35]他用数值来表示个人的愉悦和痛苦。边沁提出，经济决策是个人计算如何达到自我效用最大值的结果。[36]假如说你喜欢苹果而不喜欢香蕉，当要在苹果和香蕉中作出选择时，你会因为计算出食用苹果会给你带来更大的效用而选择它。但对于其他人来说，或许香蕉的效用更高。边沁的理论在18世纪晚期的知识分子和政治圈里吸引了一大批追随者，他的理论即后来被人所熟知的功利主义。功利主义者的信仰在于，社会应该以实现集体效用或幸福最大化的方式进行组织。

大约50年后，德国经济学家赫尔曼·海因里希·戈森（Hermann Heinrich Gossen）在边沁理论的基础上，提出了边际效用递减规律（law of diminishing marginal utility）。[37]这个规律从本质上来说是杜尔哥理论的另一面。杜尔哥指出扩大生产的收益是不断减少的，戈森则认为增加消费的收益也是不断减少的。比如，某人饥肠辘辘，于是买了一个甜甜圈，这次消费能为他带来极大的满足感。如果他仍感到饥饿，又买了第二个甜甜圈，这也能带来满足感，但是根据边际效用递减规律，第二次购买的满足感远远小于第一次。在这个人购买第五个或第六个甜甜圈时，他因购买次数增加而体验到的满足感或许会变得非常小（或者说甚至有可能是负值，因为他可能会吃到胃疼）。到某个时刻，这个人会说："我吃饱了，下一个甜甜圈不值得我花钱了。"就像农民会在价格上涨时加大生产，而在价格下降时减少生产一样，消费者说出"不值得花钱"而停止消费的时刻，会因价格的不同而到来得或早或晚。因此，价格上涨时需求会减少，反之亦然。此外，就像不断减少的收益会阻止农民无限量地扩大庄稼种植一样，不断减少的边际效用也会阻止消费者无限量地购买甜甜圈。

生产上的收益递减和消费上的边际效用递减结合在一起，意味着市场拥有一种天然的均衡机制——价格。价格是生产者与消费者共享的一个关键信息，价格上涨会同时抑制消费者的消费并鼓励生产者的生产，而价格的下降则会导致相反的结果。

为此，古典经济学以一个令人瞩目的框架取代了原来的理论，描绘

出市场是如何平衡消费者需求与经济生产，并自然而然地达到一个双方都满意的平衡点的。但仍有一个重要问题悬而未决：对于某种特定的商品、特定的效用和特定的生产过程而言，价格到底是什么？我们能够对它进行计算和预测吗？

新科学的梦想

在古典经济学家的工作之后，我们迎来了边际主义时代（1830—1930年）。这个阶段的核心人物是莱昂·瓦尔拉斯（Léon Walras），他于1834年出生于法国埃弗勒。年轻的瓦尔拉斯在事业的起步阶段步履蹒跚，丝毫看不出日后大有作为的迹象。作为一个学生，由于数学太差，他曾两次被赫赫有名的巴黎综合理工大学拒绝。后来，瓦尔拉斯转而进入国家南特高等矿业学院，却未能成为一名工程师，随后又尝试过写小说，但也以失败告终。1858年的一个晚上，郁郁不得志的瓦尔拉斯与父亲（教师兼作家）一起散步，谈到了他该如何生活。[38]瓦尔拉斯的父亲非常尊崇科学，他谈到19世纪仍有两大挑战尚未解决：即发明一套完整的历史理论和一套科学的经济学理论。他坚信可以将微分学应用到经济学上，由此创造出“经济力量的科学，类似于天文力量的科学”。[39]年轻的瓦尔拉斯受到父亲对于科学经济学愿景的启发，决心将实现此愿景作为毕生的目标。瓦尔拉斯花了数年时间，一边作为报纸作家和银行雇员艰苦工作，一边利用业余时间撰写关于经济学的文章和小册子。1870年，经过其他教授反复讨论，瓦尔拉斯最终被任命为洛桑学院的教师，并且在1872年完成了巨著《纯粹经济学要义》（*Elements of a Pure Economics*）。[40]

在瓦尔拉斯发表《纯粹经济学要义》之前，经济学并不属于数学领域。许多早期的经济学家，例如亚当·斯密和边沁，都自认为是哲学家而非科学家，而且古典时期的数学总的来说还只局限于一些数字案例和一点点几何，并没有什么复杂的东西。[41]瓦尔拉斯及其同伴一起彻底改变了这种状况。他们生活在一个科技进步的时代。在17世纪牛顿的重大发现之后，一系列的科学家和数学家，包括莱布尼茨（Leibniz）、拉格朗日（Lagrange）、欧拉（Euler）和汉密尔顿（Hamilton），开发了一种新的数学语言，他们用不同的方程式来描述越来越多的自然现象。自古希腊时期起就困扰人们的一些问题，从星球的运动到小提琴琴弦的颤动，突然之间被人类征服了。在这些理论上取得的进展让科学家非常乐观地认为他们可以用方程式描述自然界的任何方面。[42]瓦尔拉斯和他的

同胞们相信，如果微分方程式能够捕获宇宙中星球和原子的运动方式，那么数学方法也可以捕获人类大脑中关于经济学的运作方式。

瓦尔拉斯提出的经济系统中的均衡点的观点与自然界中的均衡点的观点可以类比。[43]自然界中的许多系统都有均衡点，这一概念也同样适用于物理学。正如之前章节所描述的那样，你可以想象有一个大玻璃碗，碗底呈光滑的圆弧形，你拿着一个橡皮球将其放在碗的边缘，并让它下落。球会来回滚动一段时间，但最终会在碗的底部停下，处于静止状态。静止状态的实现是由于作用于系统的所有外力互相抵消，系统实现均衡。在这个案例中，向下的重力几乎完全等于球下方的碗所起的支撑力。球将在碗底的同一位置永远保持不动，直到有其他外力作用于它。这个例子中只有一个均衡点——碗的最底部。无论我们让球往下掉多少次，它都将在同一个点停下来。

在物理系统中，均衡的类型是多种多样的。比如铅笔在笔尖立起的平衡。如果你能将铅笔立得笔直，它就能处于均衡状态。然而，与碗中的球不一样的是，这种状态是一种极不稳定的均衡，一阵微风都有可能将铅笔吹倒。此外，还有动态均衡状态。当行星在环绕恒星的轨道上运行时，恒星对于行星向内牵拉的万有引力几乎被向外推动行星的离心力冲抵了。这种均衡可以长期维持下去，使行星在稳定的轨道上运行，直到被外力打破均衡。

瓦尔拉斯将数学应用到经济学的目标之一是使经济系统具备可预测性，但不稳定的均衡从来就很难预测，这是因为小小的变化就会导致系统转向不同的方向。同样，在瓦尔拉斯的时代，判断动态系统是否处于稳定的均衡状态被认为是一个难题，为此瑞典国王奥斯卡二世承诺以2 500克朗奖金奖励解决这个问题的人。[44]如果系统拥有多个均衡点，那么预测系统将在哪一个均衡点停留就是所有几乎不可能解决的问题中难度最小的一个。瓦尔拉斯希望发掘可预测性，那就意味着他需要一个唯一的、稳定的均衡点。更确切地说，瓦尔拉斯将市场中的供求平衡比作物理均衡系统里的力量均衡。他推测，市场上每一件用来交易的商品都只有一个可以达到平衡点的价格，在这个平衡点上交易双方都很满意，市场能够出清。我们可以预测市场上的价格会在唯一的一个平衡点上稳定下来，就像我们可以预测橡皮球会在光滑的碗底静止。

为了将其推测转变为方程式，瓦尔拉斯搜遍了那个时代的物理教科书，由法国数学家路易·潘索（Louis Poinsot）于1803年发表的《静力学要义》（*Elements of Statics*），对瓦尔拉斯产生了特别重大的影响（他

甚至将自己的著作命名为《纯粹经济学要义》以致敬）。[45]正如瓦尔拉斯的传记作者威廉·贾菲（William Jaffé）指出的那样，尤其是从瓦尔拉斯提出的“用方程式表达的平衡条件”开始，瓦尔拉斯将物理学中的均衡概念引入了经济学，这为如今我们在教科书和期刊上看到的传统经济学打下了数学基础。[46]这是一个值得铭记的历史细节，因为正如我们将在下一章看到的那样，一些批评者认为，从物理学中借用均衡概念是一个重大的科学失误，它造成的影响延续至今。

为了打造均衡模型，瓦尔拉斯将在经济中占据一半地位的生产置于一旁，而将注意力全部集中在消费者之间的交易上。在他的模型中，瓦尔拉斯假定各式各样的商品已经存在于经济之中，问题在于如何设定价格以及商品是如何在相关个体中进行分配的。为了弄清瓦尔拉斯的模型是如何工作的，我们可以假设有一个大房间，里面挤满了人。每人天生就拥有一些随机的商品。比如，我拥有5个香蕉、1台洗衣机、2双鞋、5个汽车轮胎，而你有可能被分配了1条牛仔裤、2把雨伞、1部手机、3个牛油果和其他一些东西。每个人对不同商品的效用的理解不一样，例如你可能想要香蕉，而我想要牛油果，但我们都很重视手机。由于最初的商品分配是随机的，所以所有参与者对所获得的东西都很满意基本是不太可能的，因此他们想要进行交易。瓦尔拉斯认为，这种想要交易的欲望是系统失衡的标志。它意味着存在另外一种能使大家更加开心的商品分配方式，问题在于如何找到一种让房间里所有人都满意的商品分配方式，并且找到合适的商品价格，让人们可以从最初的状态转变为更加满意的状态。新的状态将会是均衡的，因为一旦所有人都尽可能满意了，鉴于可用的商品和价格只有那些，就不会再有人进行交易。[47]

为了使得交易更加有序（并且在数学上更加简单），瓦尔拉斯假定人群中有一位拍卖师。他假定其中一种商品可以作为货币（例如黄金、玻璃珠或贝壳），拍卖师会用该商品作为单位来给东西定价（比如1个牛油果值10颗玻璃珠），为每件商品编号，并记下投标价。如果买家对于某件商品的需求大于供应，他就会提高价格；反之，他就会降低价格。拍卖师会为房间中的每一件商品做这样的事，直到所有人的供求关系都达到平衡。在所有价格都确定后，人们才会进行交易，这样可以确保所有参与者都能在交易中实现价值的最大化。交易会使大家从最初随机、失衡的状态，转变为更加满意、平衡的状态。瓦尔拉斯将这种状态称为一般均衡点（general equilibrium），将他所描述的拍卖过程称为“tâtonnement”，这是一个法语单词，意思是“探索”，因为拍卖师是通过为不同商品找到不同的价格而探索出了一般均衡状态。

瓦尔拉斯的观点很新颖，但真正具有革命意义的是他将物理学中的复杂数学应用到了经济学领域。如果你能接受瓦尔拉斯的假设——人们都有不同的效用要求，他们在实现效用最大化的过程中是理性的、受自我利益驱使的，那么你就能够利用数学的精准性来预测人们将如何进行交易及如何定价。这其中还遗漏了几个小细节，比如完美的拍卖师，以及人们如何观察和衡量个人的效用这个问题，但这些问题将在未来被解决——为了实现历史上第一次对一些东西实现数值精准的、科学的预测，这只不过是一个小小的代价。[48]瓦尔拉斯希望在现实交易中实现数学上的可预测性，他的这个愿望为20世纪的经济学家树立了目标。

像万有引力一样可以预测

瓦尔拉斯并不是他那个时代唯一一个穷尽物理学教科书来寻找灵感的人。威廉·斯坦利·杰文斯（William Stanley Jevons）于1835年出生于利物浦一个富裕的工人家庭，他在11个孩子中排行第9。[49]跟瓦尔拉斯一样，他也是大器晚成。杰文斯大学肄业，在他20岁到30岁期间恰逢澳大利亚淘金热，他在悉尼的铸币厂当了一名试金员。然而，他的脑子很灵活，对铁路非常着迷，试图在业余时间里为铁路经济学建造一个数学模型。这段经历让他深信经济学可以成为一门数学科学。后来杰文斯决定返回英国，学习经济学，取得学位。同瓦尔拉斯一样，杰文斯肩负着一种使命——“定义人类知识的基础”，以及“在合理的基础上重塑（经济学）科学”。[50]

1867年，两位卓越的英国科学家威廉·汤姆森爵士（Sir William Thomson）及皮特·格思里·泰特（Peter Guthrie Tait）合著了一本名为《自然哲学论》（*Treatise on Natural Philosophy*）的教科书，它集合了当时能量物理学方面的最新发现。[51]杰文斯就是这本书的狂热读者之一。在汤姆森和泰特的书中，杰文斯发现了由迈克尔·法拉第（Michael Faraday）和詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）发明的用于描述万有引力、磁力和电力的新理论——“力场”。比如，一个质量块（如太阳）会产生一个引力场，将其他物品吸引到它身边。质量块越大，它的引力场越强。杰文斯认为人的利己主义就像一个引力场：[52]“效用只有在一个人想要某样东西，而某样东西被他需要的时候才存在……就好像某种物质所产生的引力不仅取决于物质本身的质量，还取决于周边物质的质量以及双方的相对位置和距离，因此，效用是指希望得到某物的人以及被需要的某物之间的吸引力。”[53]

杰文斯在1871年出版的《政治经济学理论》（*Theory of Political Economy*）一书中沿用了边沁提出的效用概念以及戈森提出的边际收益递减理论，并使用从场理论中进化而来的方程式来将他们的理念从哲学概念转化为数学模型。[54]简而言之，杰文斯希望能使人类行为像万有引力一样可以预测。为了预测物体在引力场中的运动方式，有两件事是必须知道的：引力作用的方向以及物体运动过程中受到的阻力的方向。回到之前的案例，如果我们将球放入碗里，重力会使球向下运动，而碗的内壁会阻止球的运动。如果我们知道“向下”是指什么方向以及碗的阻力的方向，就能预测出球将停留在何处。同样，如果我们知道引力的吸引方向以及阻碍运动的钟摆长度，我们就能预测出钟摆的均衡点。在杰文斯的概念中，利己主义提供了一种类似于万有引力的力量，它会将我们吸引到幸福或效用最大化的地方。但人们也生活在一个资源有限的世界里，这为我们的行为增添了阻力。窍门在于找到一种商品和服务的组合，能够在有限资源的阻碍下，实现幸福最大化，就像瓦尔拉斯提出的模型一样，我们通过交易来实现这种状态。

假设经济中只有两种商品：酒和奶酪。通常情况下，我喜欢酒胜过奶酪，但根据边际效用递减规律，在某个时刻我会觉得自己拥有的酒足够多了，那时我更希望要一些奶酪而不是一杯酒。相反，你通常喜欢奶酪胜过酒，但在某个时刻，你会更愿意要酒而非奶酪。我们假设，市场中的酒和奶酪的数量都是有限的，我们随机分配到了这两样东西。跟瓦尔拉斯的模型一样，我们不可能正好获得与自身效用相符的数量，于是我们就会进行交易，直到每个人都获得数量最满意的酒和奶酪。

杰文斯的持久贡献在于将经济学的选择问题描述为有限条件下实现最优化的练习。也就是说，在可用数量的前提下，消费者要计算出能够实现个人最大幸福的不同商品的数量。在杰文斯看来，个人效用的差异创造出了一种潜在的交易能量。他在《经济学原理》（*Principles of Economics*）中写到：“价值的概念之于科学，就相当于能量的概念之于数学。”[55]就像碗中的球在碗壁的阻碍下寻求最低的能量状态一样，人类也会在有限资源的限制下寻求最大的幸福状态，并通过交易的方式实现这个目标。

趋于至善的经济

亚当·斯密假定人类的利己主义会驱使市场走向平衡，在这种平衡状

态下，人们会商定价格，进行交易，出清市场。瓦尔拉斯证明了这种平衡状态可被认为是一个平衡点，并且可以用数学计算出来。杰文斯表明，在效用不同、资源有限的世界里，如果人们追求幸福的最大化，他们必然要进行交易来达到市场的平衡点。亚当·斯密推进了杰文斯的观点：利己主义不仅能使市场趋于平衡，它还会为整个社会带来最好的结果。

意大利人维尔弗雷多·帕累托（Vilfredo Pareto）是跟瓦尔拉斯和杰文斯同时代的人。帕累托接受过工程师的训练，并写过题为《固体的弹性平衡》的博士论文，他同瓦尔拉斯和杰文斯一样精通物理学，甚至更胜于他们。[56]帕累托是一个性情古怪的人，晚年同20只安哥拉猫一起隐居在瑞士山中的小屋里。他也因发明了经济学领域的重要概念之一而在经济学领域声名大噪。

在亚当·斯密的《国富论》发表之后，经济学家希望能够断定竞争市场是否真的实现了社会利益的最大化，若真如此，又是在何种情况下实现的。尽管杰文斯已经极大地推进了效用的理论研究，但仍存在一个问题，即效用的不可衡量性——我们不可能看透人们的脑袋，计算他们的效用，然后把它们相加。那么，如何才能判断社会利益是真的提升或者实现最大化了呢？

帕累托通过一个天才的逻辑推理解决了这个问题。他推断说，人们可以进行四类交易，第一类是双赢的交易，双方都可以获利。在这种情况下，很明显社会利益可能得到提升；第二类是只有一方获利，但另一方没有损失，社会利益毫无疑问会再次得到提升；第三类是没有人获利，但有人失利的交易，这种情况下社会利益毫无疑问将会下降；第四类是有人获利，也有人失利的交易，但由于无法直接计算效用，所以也就无法判断净影响。帕累托认为，既然是同意交易的双方参与其中，他们又不傻，所以人们只会选择双赢或一方获利而另一方无损失的交易，这两类交易都会提升参与者的总利益，这些交易后来被称为帕累托改进交易（Poreto superior trades）。帕累托主张，在自由市场中，人们会不断进行交易，直到完成所有帕累托改进交易。在那个时刻，交易将会停止，因为进一步的交易将会导致某些人的利益受损，而市场将由此达到平衡点，这就是后来经济学家所谓的帕累托最优（Pareto optimal）。因此，帕累托最优是指如果不损害任何人的利益，就不可能再继续进行交易的点。帕累托最优未必是全体价值最大化的点，因为有些交易可以在损害部分人利益的情况下为其他人带去好处，从而提升整

体的总效用。由于无法准确地对效用进行计算，也没有独裁者可以为了某些人的利益而强制损害他人的利益，因此，帕累托最优是自由社会所能达成的最好结果。[57]

根据瓦尔拉斯、杰文斯、帕累托以及其他边际主义者的理论，在市场经济中，参与者会自由交易，达到可用资源条件下尽可能满意的状态。通过这样的交易，经济会日渐趋于平衡，即达到一个自然的静止点。在该状态下，供求相等，资源的利用效率最高，社会利益达到帕累托最优。用伏尔泰笔下的人物潘格洛斯博士的话来说："在理想的世界里……一切都是为了最美好的目的而设。"[58]边际主义者最突出的成就或许在于让经济学拥有了一种数学理论，这种理论表明，如果顺其自然，自由市场经济将会达到一种潘格洛斯状态（趋于至善），这就像球在碗底静止一样是不可避免的。瓦尔拉斯声称他的"经济学的纯理论是一门从各方面来说都与物理、数学相似的科学"。杰文斯相信自己发明了一种"具有道德影响力的微积分学"。帕累托则声称："经济科学的理论由此养成了理性力学的严谨。"[59]在他们看来，边际主义者成功地将经济学转变成了数学科学。

新古典综合派

20世纪，在边际主义者奠定的基础上，伟大经济学家的先贤祠整合建立起来了。在世纪之交，英国经济学家阿尔弗雷德·马歇尔将杰文斯关于单个独立市场的模型（部分均衡）与瓦尔拉斯关于多个关联市场的模型（一般均衡）联系在了一起。马歇尔首创了交叉供求曲线图，此后，这个曲线图就一直在给经济学系的学生们带来难题。20世纪30年代，约翰·希克斯（John Hicks，他曾在曼彻斯特大学担任杰文斯的专属教授）在其作品《价值与资本》（*Value and Capital*）中将瓦尔拉斯、马歇尔和帕累托的成果综合为一套连贯的理论。20世纪中期，欧洲陷入战争，创新的中心转移到了大西洋彼岸，在那里，一代美国人以及由希特勒制造的欧洲难民创造了经济学理论的现代核心，即后来的"新古典综合派"。保罗·萨缪尔森（Paul Samuelson）和肯尼斯·阿罗（Kenneth Arrow）是那个时代两位最为卓越的人物。

萨缪尔森是一个名副其实的神童。[60] 1941年，他26岁，他在当年完

成了《经济分析基础》（*Foundations of Economic Analysis*），这是萨缪尔森在哈佛大学读研究生时完成的一篇论文。在这篇文章中，他主要采用了希克斯的综合理论，并加上自己的创新，将之转化为一个引人瞩目的数学理论，这后来成了市场运作的标准模型。[61]萨缪尔森的突破性贡献在于解决了一个自边沁时期起就困扰着经济学家的问题。当时效用已经成为经济学理论的一个关键部分，但它仍是一个不可观测、不可度量的神秘变量。帕累托和希克斯提出，“尤特尔”是计算效用的固定单位（就像千克或瓦特），并且认为效用只有在相对情形中才有意义，就好比“对我而言，那个苹果的效用是橘子的两倍”。但这依然引出了一个问题，即如何计算相对效用。萨缪尔森的回答是，我们不需要窥探他人的脑袋而直接计算效用，人们作出的选择会揭示他们的意愿。我们需要做的就是假定人们在行为中会表现出逻辑和一致性。比如，你让某人在苹果和橘子中作出选择，他选择了苹果，那么就可以预测如果要在苹果、橘子和香蕉中作选择，他不会选择橘子。按照逻辑来说，他应该依然喜欢苹果胜过橘子，所以他会选择苹果或者香蕉。尽管这个结果无法让我们得出“一个苹果的尤特尔是橘子尤特尔的两倍”的结论，但是我们可以肯定地说，在这种情况下，问题中的主人公“喜欢苹果胜过橘子”。萨缪尔森认为，这个简单的论断足以建立一个需求侧的理论，因此他用一套关于人们偏好排序的、基础的、逻辑学的规则替代了效用理论。这些规则后来成了传统经济学消费行为理论的基础，以及“人们在作经济选择时是理性的”这一理念的主要支撑。[62]

比萨缪尔森年纪稍小的肯尼斯·阿罗同样在很小的时候就展现出了非凡的数学才能。阿罗为经济学领域作出了许多基础性的贡献，其中最著名的是他同法国经济学家吉拉德·德布鲁（Gérard Debreu）在1954年证明的理论。阿罗和德布鲁将瓦尔拉斯关于一般均衡的理念与帕累托关于最优的理念用一种一般性的方法联系起来，从而创造了关于一般均衡的新古典理论。他们的理论表明，经济体中的所有市场都能够自动协调出一套价格，那就是经济整体的帕累托最优，这种现象哪怕在市场存在不稳定性的时候也会发生（瓦尔拉斯在他的模型中规定一切都是确定的）。[63]这种自动协调的产生是由于市场之间是相互联系的——某些产品可以成为其他产品的替代品，比如咖啡的价格上涨，消费者可以转而选择茶，以及有些商品是与其他商品一起消费的互补品（例如，汽油价格的上涨有可能会降低人们对于高耗油的大型汽车的需求）。阿罗和德布鲁表明，价格就像神经系统，负责在整个经济体中传递供应和需求的信号。信奉利己主义的人们会根据信号作出反应，不可避免地驱使系统

达到社会最优的均衡点——“看不见的手”真的是威力十足。

阿罗-德布鲁一般均衡理论最惊人的成就或许就在于，这个极其有力的结果是基于少数几个公理建立起来的。其中的部分假设毫无争议，例如你不能进行负劳动或负消费。然而，其中的另一部分假设则存在更多令人质疑的地方。例如，该理论假设，每个人都或多或少拥有每一种商品，未来市场里存在着每一种商品和服务，每个人在作计算决策的时候都极其理性并且了解世界未来可能出现的所有状况。在瓦尔拉斯最初的模型中，这些假设被当作必要的简化条件，细节要另行讨论。重要的是，我们可以以一套简单的公理作为起始点，然后严格地用数学方法得出一套通用的结果：在竞争市场中运作的理性利己主义将驱使经济达到最优点。1954年，在这套理论发表的时候，它被经济学家认为是一种重大突破。在冷战高峰时期，它最终在政治领域被误读成了资本主义市场优于社会主义市场的证据。[64]可以肯定的是，阿罗和德布鲁的理论是经济高度简化的影像，它缺乏真实世界的特征，例如垄断性产业、工人联合会、政府监管、税收等，但它的政治信息却很清晰。在没有扭曲和干扰的情况下，我们越接近完美市场竞争的理性状态，就越接近最优均衡。

20世纪60年代出现了一种基本完整的理论，这种理论从关于个人消费者和生产者的公理假设出发，呈现了关于市场和经济学的全面结论。经济学家将这些自下而上的、关于个人和市场的理论称为微观经济学。在此期间，关于宏观经济学的研究也有很多。在宏观经济学中，经济学家采用自上而下的视角提出了一些问题，例如为何存在失业，是什么导致了商业的循环以及利率与通货膨胀之间存在怎样的关联。我们将在后续章节讨论这些内容，但当前的关键问题在于，在20世纪60年代至70年代，芝加哥的经济学家（这样说是因为他们大部分人都是芝加哥大学的教员），例如米尔顿·弗里德曼（Milton Friedman）和罗伯特·卢卡斯（Robert Lucas）开始将新古典微观经济学的理论应用到宏观经济学上，诸如理性效用最大化的消费者、最优均衡之类的概念也成了传统宏观经济学理论的核心部分。

从分配到增长

在本章的前半部分，我曾提到过历史上经济学曾被两大问题困扰：一是财富是如何被创造出来的，二是财富是如何进行分配的。在亚当·斯

密的古典时期以及萨缪尔森和阿罗的19世纪中叶，第一个问题被第二个问题掩盖了。瓦尔拉斯、杰维斯以及帕累托的模型都是从一些假定条件开始的，即经济已经存在，生产者拥有资源，消费者拥有各式各样的商品。在这些模型里，关键问题变成了如何通过为每个人提供最大利益的方式，分配已经存在的有限资源。为何他们的注意力都集中在对有限资源的分配上？这里边有一个重要的原因，那就是从物理学中引用的数学均衡方程式在回答分配问题上是极为理想的，但要想将它运用到增长方面会十分困难。在定义上来说，均衡系统是一种静止的状态，而增长则意味着变化和动态。

意识到均衡与增长之间存在矛盾的是约瑟夫·熊彼特，他通常被说成是奥地利经济学家，但他其实出生于今天的捷克共和国。[65]熊彼特因为穿着马靴去哈佛大学教师会议、穿着正式的晚礼服出席家里的晚餐而变得人人皆知，他还因宣传自己的人生三大目标而出名：成为维也纳最伟大的情人、欧洲最伟大的马术师以及世界上最伟大的经济学家。但他或许会说，可惜自己没能达成第二个目标。在财富的分配问题上，熊彼特认同新古典主义同侪的均衡概念，但他不相信这是回答增长问题的正确框架。新古典主义关于生产的观点是静态的。他们假定公司拥有固定的技术和产品组合，人们所需要做的就是计算出能够实现利益最大化的产品数量。但在熊彼特看来，经济增长不仅是指增加已有产品的数量，还与创新相关："将车厢连续相加，加多少都可以，但你永远也不会因此而得到一辆货车。"[66]用第1章中给出的术语来说，熊彼特希望解释的是库存单位增长以及数量增加。

新古典主义者倾向于将创新看作外部的、外生性的因素，认为它是影响经济的一种随机变量——就像天气，认为它不在经济学的研究范围内。然而，熊彼特却认为，我们必须将创新看作一种经济内部的、内因性的因素，并且这对于理解经济是非常关键的。他坚持主张，为了实现增长，必须由"经济系统内部的能量来源打破将要达到的均衡状态"。[67]对于熊彼特而言，这种能量的来源就是创业者，他几乎是用英雄主义的笔触来描写他们的。在熊彼特看来，技术进步因一系列随机的发明而产生。然而，新技术的商业化却面临许多障碍，包括从融资需求到旧习惯和旧思想的阻碍。因此，就像水坝后面的水一样，发明的随机雨水会逐渐积蓄起来。在熊彼特的理论中，创业者扮演着大坝破坏者的角色，是他们将创新之洪泻入市场的。通过这种方式，经济就会出现增长，但这并不像平稳的溪流，用熊彼特的话来说，这是"破坏性的创新风暴"。根据熊彼特的观点，财富的起源归功于各位创业者历史性的努力。在熊彼

特的理论里，财富的起源是经过理查德·阿克赖特、亨利·福特、爱迪生和乔布斯这样的人战胜种种困难，将时代的技术转化为成功的商业企业而产生的。

熊彼特的理论在本质上是关于人类和历史的理论，它既有优势，也有不足。尽管熊彼特理论的言辞丰富性使得它在今天仍能引起共鸣，但他从未成功地将其转变成严谨的数学语言。这就意味着，他的理念永远不可能与数学式的新古典主义的框架达成和解——这是最终限制了熊彼特理念影响力的不足之处。[68]缺乏数学方法也使得增长理论在随后的40年里始终停滞不前，直到罗伯特·索洛（Robert Solow）的出现。[69]

索洛出生于布鲁克林，他在哈佛大学接受训练，并在麻省理工学院度过了自己的职业生涯。[70]索洛并没有被熊彼特的缺乏数学敏锐性所困扰，他试图将增长与新古典主义理论中所说的碗中球的可预测性调和在一起。在1987年获得诺贝尔经济学奖并发表演讲时，索洛描述了他发展理论的动机。[71]关于增长的早期数学工作被过度简单化了，它假定资本的产出是恒定不变的，也就是说，对于工具、机器和装备等进行投资所得到的回报是恒定不变的。这种假设明显是不现实的。从历史发展来看，技术的变革将会极大地提升资本的产出——拖拉机的产出要远远高于牛拉犁的产出。索洛希望找到一种方法来对这个重要的方面进行解读。熊彼特将创新看作打破均衡的力量，而索洛则希望用一种与新古典主义理论一致的方式来解读创新，并维持经济的均衡。

增长和均衡听起来并不像一对可以兼容的概念，碗中球所处的系统并不是增长的。然而，在1956年发表的一篇标志性论文中，索洛将两个概念协调在了一起——索洛将经济视为一种动态均衡，或者说是他所谓的平衡增长（balanced growth）。[72]请想象有一场马戏表演，一位勇敢的演员在高空钢丝上踩单车。为了保持平衡并且不从钢丝上掉下去，演员手里会拿一根长长的横向木杆。虽然演员在钢丝上骑车是运动着的，但在每一个时间点上，这位骑手都处于某种平衡状态。同样，索洛将经济看作均衡中的平衡，哪怕它处于不断增长中。他在模型中提出有两个关键的变量是外生性的：人口增长率和技术革新率，这两个变量驱动了经济增长（你可以将它们看作高空钢丝骑手作用于单车踏板的力量）。随后，索洛表明，经济中的其他因素，例如储蓄率和经济中的总资本量，将根据人口增长和技术变革自动地进行平衡，就像马戏演员不断地移动木杆来保持平衡。在索洛的模型里，平衡单车骑手的角色是由劳动力和资本市场扮演的，即便是在经济不断增长的情况下，它们也在努力

地使一切维持在帕累托最优均衡的状态。

索洛的模型与亚当·斯密的某些观点是一致的，即尽管人口增长或许能够增加国家的总财富，但只有提高生产力才能提高国家的人均财富——一个国家的富有不在于它拥有多少资本，而在于这些资本的生产力有多高。在索洛看来，生产力的关键在于技术。索洛的模式意味着，美国和其他西方国家并不是因为拥有自然资源，或者拥有天上掉下的资本而变得富有的。它们变得富有的过程是一个良性循环，即技术的改进提升了资本的生产力，这又会提高储蓄水平，继而增加投资资本。如果没有技术的提升，资本只会根据人口数量的情况按比例增长，而人均财富值只会保持不变。1956年距离“知识经济”一词变得流行还有很久，但罗伯特·索洛已经发现了它。[73]

索洛的工作使得人们对于增长这个主题重新燃起了兴趣。很快，在他提出的基本模型的基础上又出现了一系列变化的版本。然而，在20世纪80年代，以斯坦福经济学家保罗·罗默（Paul Romer）为首的一群研究者越来越不满意索洛模型提出的“增长的真正驱动者——技术是外生性的”这一说法，这就跟熊彼特在50年前受到的挫折一样——当时经济学家认为创新也是外生性的。[74]罗默认为增长的“能量”应该被看作经济的内在因素，1990年，罗默发表了一篇论文，启动了后来被称为“内生增长理论”的研究。[75]

索洛没有将增长的能量源泉定义为创业者的英雄主义，而是定义为技术的特性。他指出，技术对于增加具有一种累计的、加速的特性。我们知道得越多，现存的人类知识的基础就越大，并且从下一个发明中得到的利益就越多。知识被经济学家称作一种沉浸回报现象。正如之前讨论的那样，18世纪，雅克·杜尔哥指出，大部分生产过程表现出了收益递减的相反特性。对于大多数的生产过程来说，无论是农业、制造业还是服务业，随着投入资源的不断增加，边际收益会越来越少。罗默认为，在生产技术的情况中（比如我们可以将研发当作生产技术的一个环节），这个逻辑是相反的。如果我们在知识上的投入随着时间的增加而不断累积，收益就会越来越高。今天，人们在微芯片和生物技术上投入一小时进行研发所得到的收益，比1900年在蒸汽机和电报机上投入一小时进行研发所得到的收益要高。罗默在他的模型中创建了一种积极的反馈、一种良性的循环，在这种循环里，在技术上投入得越多，一个社会就会越富有，未来在技术上的投资收益也就越高，结果就是无限的、指数级的增长。在技术投资上不断增加的回报将赋予经济增长的脚踏板以

前所未有的速度和力量。

传统经济学的遗产

到20世纪末期，传统经济学完全由新古典主义的思维框架主导，其基础理念就在于，理性、乐观的消费者和生产者在资源有限的世界里做决策，并且（在没有技术投资的情况下）这些决策所带来的收益是递减的。利己主义和种种限制会驱使经济达到帕累托最优均衡。经济分析的方法论同样被数学证明主导着，这些数学证明都以一套假设作为前提，并且有逻辑地构建了一套结论。以索洛为先驱的新古典主义增长理论声称回答了财富的起源这一伟大问题，而阿罗和德布鲁的新古典主义一般均衡理论从表面看来回答了财富的分配这一伟大问题。当然，这些权威的模型有许多不同版本，包括以不确定性、不完美竞争、不完全竞争为特征的模型。但这些模型针对的是同一个主题的不同版本，而不是新的乐章。

20世纪的经济学家意识到，他们雄心勃勃地想要创造一套严谨的、定义明确的数学模型来描述经济的运作。尽管在新古典主义的框架下，将微观和宏观视野全面融合的梦想没有完全实现，但人们仍然可以通过一个逻辑连贯的框架和一套假设，历经从个人决策者的原子世界到饱览国家经济的旅程。[76]

毋庸置疑，传统的思维框架对公共政策、商业和金融世界产生了重大影响。从中央银行行长到总统顾问、财政部部长，这些政策制定者全都依赖于传统经济学的概念和模式。同样，传统经济学理论中的概念通常被用于商业领域的决策——决策涵盖了从竞争策略到是否进行并购或收购。[77]此外，毫不夸张地说，通过传统经济学理论计算的世界金融市场的交易量每天高达数万亿美元。传统经济学的理念在促进我们对经济和社会的广泛了解方面作出了巨大的贡献。

尽管传统经济学起到了毋庸置疑的巨大作用，但本章伊始提出的焦虑仍未解决。经济学家沃纳·希尔登布兰德（Werner Hildenbrand）曾将一般均衡理论比作一座哥特式大教堂，瓦尔拉斯及其同辈经济学家都是这座教堂的建筑师，而20世纪的伟大经济学家们则是工程队，但这座教堂的地基很不稳定。[78]

03

一场动摇传统经济学根基的辩论

虽然约翰·里德对经济学家很失望，但他依然与这个群体保持着积极的联系，并在久负盛名的拉塞尔·赛奇基金董事会任职，该基金会旨在支持社会科学研究。有一次，在纽约董事会的茶歇时间，另外一位董事、史密森学会秘书鲍勃·亚当斯（Bob Adams）告诉里德，一个全新的研究机构正在新墨西哥的沙漠上崛起。[1]研究团队的领头人是原白宫科学顾问、洛斯阿拉莫斯国家实验室的科研负责人乔治·考恩（George Cowan）。他的合伙人还包括一批科学界最耀眼的明星，包括诺贝尔奖得主、夸克的发现者默里·盖尔-曼（Murray Gell-Mann），来自普林斯顿大学的诺贝尔奖得主菲利普·安德森（Philip Anderson）以及数位来自洛斯阿拉莫斯国家实验室的资深研究员。

这个团体设立了一个雄伟的目标，即从根本上改变科学研究的方法。在历史上，科学采取了一种自上而下、归纳式的方法，将宇宙分割成许多小块，在星系到亚原子粒子之间寻找终极法则。圣塔菲研究所的科学家们认为，尽管这种方法取得了巨大的成功，但许多难解的问题本身就是“复杂系统”，它们具有集体或涌现的特征，以自下而上、系统的方法来理解会更好。[2]例如，他们认为，类似于“什么是生命”这样的问题，是永远不可能通过从有机化学方面自上而下的审视来进行分解的。[3]有机体是一个复杂系统，其整体要比各种化学部件的总和更加复杂。

要想回答“什么是生命”这个问题，需要一种将有机体视为整体的视角，以及自下而上的、对于数十亿分子如何互动而创造出一支复杂舞蹈（即所谓的生命）的理解。该团队认为，对于一系列广泛的现象，包括大脑、生物生态系统、互联网以及人类社会本身，这些现象的整体要比组成部分庞大得多，因而需要采用上述方法。他们还相信，这些难解的科学问题需要多学科的视角。“什么是生命”这个问题的解答需要生物学家、物理学家、化学家、计算机科学家以及其他科学家共同努力。然而，大部分大学和研究实验室都是以单独的院系进行组织的，这样不利于合作。1984年，该团队创建了圣塔菲研究所，这是一个非营利性组织，旨在开展对复杂系统的跨学科研究。不久之后，他们又在一处废弃

的修道院建起了办公室，考恩在修道院院长的办公室安置了下来，基督圣血山脉的景色激励着他们。[4]

里德被亚当斯关于圣塔菲研究所的描述吸引了。理解像全球经济一样复杂的系统无疑是一个困难的科学问题，或许圣塔菲的研究者提出的自下而上、多学科交叉的方法可以为经济学提供一种必需的智力上的驳诘。亚当斯向圣塔菲研究所的创始团队引荐了里德，1987年，里德和花旗银行同意资助一个关于经济学的多学科交叉工作坊。

巨人的冲突

那个工作坊的会议开得就像一场橄榄球比赛。[5]参与对决的一方是10位著名的经济学家，他们以诺贝尔奖得主肯尼斯·阿罗为首。另一方则由10位物理学家、生物学家和计算机科学家组成，他们以菲利普·安德森为领头人。经济学家一方拥有数位大咖，例如拉里·萨默斯（Larry Summers），他后来担任了美国财政部部长以及哈佛大学校长；还有来自斯坦福大学的布莱恩·阿瑟，他的理论为微软反垄断官司提供了关键论点；何塞·施可曼（José Scheinkman），他后来成了颇具传奇色彩的芝加哥大学经济学系主任。科学家队伍在吸引眼球方面一点儿也不逊色，其中包括大卫·吕埃勒（David Ruelle），他是混沌理论的先锋之一；约翰·霍兰德（John Holland），他是人工智能方面的研究者；斯图尔特·考夫曼（Stuart Kauffman），他是来自宾夕法尼亚州的生物学家，曾获得麦克阿瑟基金“天才奖”(4)；多因·法默（Doyne Farmer），他是来自洛斯阿拉莫斯国家实验室的一位年轻物理学大咖，因为使用非线性物理理论在阿拉斯加赢了轮盘赌而声名大噪。

双方都介绍了各自领域当前的最新成果，然后用了10天时间就经济行为、技术创新、商业循环以及资本市场的运作进行了辩论。经济学家们因为物理科学家们的理念和技巧而激动不已，同时也认为这些科学家在经济学问题上非常幼稚，甚至有一点狂傲。另一方面，物理科学家们则对经济学家们在数学方面的精湛技能留下了深刻印象，并且切实感受到了经济学问题的难度。

但真正让物理科学家震惊的是，在他们眼里，经济学家像是来自过去的人。后来，其中一位参会者评论说，这些经济学家让他想起了最近一次的古巴旅行。他描述说，古巴是一个因为美国贸易禁运而几乎完全与西方世界隔绝了50多年的地方。大街上到处都是20世纪50年代的帕卡

德汽车和德索托汽车，还有少数是更新一点的古董车。他指出，我们不得不承认古巴人的心灵手巧，他们利用苏联拖拉机的翻修零部件让这些汽车运行了这么长时间。对于物理学家而言，他们在经济学家身上看到的大多也是一种类似的“复古”的感觉。在他们看来，经济学家好似被封锁在了自己的知识禁运令中，与数十年的科学进步相隔绝，但同时他们也在心灵手巧地弯折、拉伸和更新自己的理论，以便让它们保持效用。物理学家们看到的是瓦尔拉斯和杰文斯遗留下来的内容，这些数学“帕卡德”和“德索托”是边际主义从100年前的物理教科书里掠夺来的方程式和技术。

不仅是经济学里的数学部分看起来像来自过去的冲击波，物理学家们还为经济学家在模型里继续使用简化假设而感到震惊。自伽利略的时代起，科学家们就开始使用简化模型，例如标准球体以及理想气体，以便让模型更容易分析。但科学家们往往会很谨慎地去确认，尽管他们的假设是现实的简化版，但简化不会与现实发生真正的冲突。科学家们还会仔细地进行测试，验证他们的假设是否与理论所给出的答案有关。以参加工作坊的科学家们的观点来看，经济学家对于假设的使用到了一种极端的程度。其中一个假设获得了科学家的特别关注，那就是经济学家所谓的完美理性。传统经济学简化了人类的行为，假设人们会千方百计地了解未来，并通过不可思议的复杂计算对所有的信息进行处理，最终作出是否购买一品脱牛奶这样的基本决策。科学家们没有意识到，对于这个问题的争辩已有很长一段历史，他们仍在强烈反对使用一个与现实不符的模型。科学作家米奇·沃尔德罗普（Mitch Waldrop）引用了其中一位经济学家布莱恩·阿瑟关于交换的描述：

> 物理学家们对经济学家们所作的假设感到震惊——模型不是现实的反面，而是为了看看这些假设是否已经成为该领域的通用货币。我只看到菲利普·安德森面带微笑地向后靠了靠说：“你们真的相信这句话吗？”
>
> 被逼到角落的经济学家会这样回答：“是的，这能让我们解决问题。如果不作假设，你就什么也干不了。”
>
> 物理学家会单刀直入：“但这些假设能起什么作用呢？如果假设不是现实，那么你所解决的不过是一个错误的问题。”[6]

在这一章节中，我将论证，尽管经济学领域取得了许多成功，但圣塔菲研究所的科学家们所提出的担忧是合理的。瓦尔拉斯将均衡的概念从物理学引入经济学，获得了数学的精准性与科学的可预测性。但他也为这种收获付出了高昂的代价——现实主义。均衡理论的数学要求瓦尔拉斯和后来的经济学家作出一系列限制性很强的假设，这使得理论经济学与现实世界的距离越来越远。传统经济学家遇到了计算机程序员所谓的“垃圾进、垃圾出、无用输出”的问题。如果你向计算机输入无用的东西，它将通过绝对的精准和完美的逻辑产出废料。同样，大多数传统经济学模型以不现实的假设作为开端，随后伴随着数学的必然性，得到的只是同样不现实的结论。正如我们将会看到的那样，这就是传统经济学的许多核心理念都很难找到实证支撑的原因。我们将检视这些让圣塔菲研究所的科学家们感到烦恼的假设，然后检视传统经济学的实证记录。在结尾时，我们将回顾经济学的历史，看看一个历史偶然事件是如何使传统经济学坠入长达一个世纪的错误低谷的。

不现实的假设

圣塔菲会议并不是经济学家和科学家就假设的使用而第一次发生冲突。1901年，莱昂·瓦尔拉斯送给具有传奇色彩的法国数学家亨利·庞加莱（Henri Poincaré）一本自己的著作《纯粹经济学要义》，并征询他的看法。庞加莱回复说：“就先验而言，我并不反对将数学应用到经济科学上，只要这不超越一定的界限。”在后来的一封信中，庞加莱明确指出，所谓的界限是指瓦尔拉斯理论中包含的许多“任意函数”（arbitrary functions，指瓦尔拉斯对于假设的使用）。庞加莱评论说，从瓦尔拉斯的方程式中得出的结论从数学的角度来说是正确的，但“如果任意函数在结果中复现”，理论所得出的结果就“没有益处”。[7]跟一个世纪之后的圣塔菲研究所的科学家们一样，庞加莱尤其关注瓦尔拉斯对于经济因素的无限制预测的假设。用庞加莱的话来说就是：“你将人类看作非常自私而又非常有远见的。第一个假设或许能被第一近似值验证，但第二个假设则需要抱持保留态度。”[8]

经济学家和当时的先锋科学家之间的情况也与此类似，这些科学家包括法国物理学家约瑟夫·伯兰特（Joseph Bertrand）、赫尔曼·劳伦（Hermann Laurent），美国先锋热力学家威德拉·吉布斯（J. Willard Gibbs），伟大的意大利数学家维多·沃尔泰拉（Vito Volterra）。他们都回应了庞加莱的抱怨，即尽管经济学变得更加精准和严谨是值得称赞

的，但为了让方程式变得易解而抛弃现实是不可取的。[9]然而，经济学家很大程度上忽略了这些批评，打造经济学的新古典主义理论的进程还在继续。关乎假设的争议并没有消散，数十年来一直在该领域持续存在。

1953年，来自芝加哥大学的米尔顿·弗里德曼发表了一篇名为《积极经济学的方法论》（*The Methodology of Positive Economics*）的论文，又将争议推至完全沸腾的状态。[10]论文指出，只要能够作出正确的预测，经济学理论中不现实的假设无关紧要。如果经济在“好似”人们完全理性的情况下运行，那么人们是否真的完全理性是无关紧要的。只要结果正确，假设无须进一步证明。也就是说，如果不是“垃圾出”，那么就不用去管“输入”的是什么。这篇论文被广泛阅读并立即引发了争议。[11]在数年之后的一次美国经济学会的会议上，来自卡内基·梅隆大学的赫伯特·西蒙（Herbert Simon）发表了相反的观点。[12]西蒙指出，科学理论的目的不在于进行预测，而在于解释事物——预测是关于解释是否正确的测试。但我们必须测试解释的整个逻辑链条，而不仅仅是末端的结论。

我将举一个简单的例子来说明西蒙的观点。你可以提出一个理论来解释天空为什么是蓝色的——假设有一个巨人，每天晚上都会在我们睡觉的时候粉刷天空。[13]作为一种极端情形，弗里德曼的逻辑会认为，只要理论能够作出正确的预测——预测天空是蓝色的，那么关于巨人的假设其实无关紧要。西蒙则认为，我们不能仅仅检测结论正确与否。要接受这样的理论，我们必须观察巨人的行动。正如经济哲学家丹尼尔·豪斯曼（Daniel Hausman）所言，我们必须“深入了解”一个理论，以确保解释的因果链条也是合理的。[14]

那么，假设在理论中的应有作用是什么呢？为何伽利略和牛顿能够假设完美真空和理想球体，而瓦尔拉斯不能假设完全理性的人和像神一样的拍卖师呢？科学哲学家认为，在使用假设时通常有两条黄金法则。[15]第一，假设必须与模型的目标相适宜；第二，假设不能影响模型为目标提供的答案。这两条规则的来源可以归结为科学哲学家所谓的细粒度与粗粒度。

要想阐释这种来源，有一个好办法，那就是将科学理论想象为地图。[16]地图是潜在现实的近似图画，艾奥瓦州奥斯卡卢萨的地图只是现实奥斯卡卢萨的相近呈现。奥斯卡卢萨唯一完美的地图是奥斯卡卢萨本身，但它太大了，无法放进汽车的仪表储物箱里，因而用处不大。就像地图制作者会将地形进行理想化或省略某些特征，科学家也会对他们的理论进行简化和理想化。包含或省略哪些内容取决于地图或理论的目

的。如果你要驾车周游全国，你或许只需要一张粗粒度的地图来为你指示主要的高速公路。但是，如果你想去拜访住在奥斯卡卢萨福特大道的姑姥姥，你就需要一张细粒度的地图，其中标注着奥斯卡卢萨的街道网络，而不是全国所有的高速公路。同样，宇宙学家或许是从星系的层面来看待宇宙的，而化学家有可能是从原子的层面来看宇宙的，每一位研究者都需要不同类型和不同程度的理想化。关键在于，粗粒度和细粒度的地图（及理论）必须相互兼容并符合潜在的现实。

如果一张高速公路地图标识某个位置有一条河，那么这条河在当地地图上必须处于相同的位置，并且必须符合现实中它所处的位置。与此类似，虽然宇宙学家和化学家的关注点不一样，但他们的模型不应互相矛盾，而应该都与经验证据和实验证据相一致。在制作地图时，我们不能为了容易绘制而移动道路和河流的位置。对于许多批评者而言，传统经济学的假设看起来并不像粗粒度的合理方案。看起来，从瓦尔拉斯和杰文斯开始，经济学家就在随意设定关于完全理性的假设、神一样的拍卖师等，而它的唯一目的就是让均衡数学能够奏效。我们现在来具体看一下传统经济学中的一些令人不安的假设。

简单到不可思议的世界里的聪明到不可思议的人

在传统经济学的所有假设中，最强有力、最不现实的或许要数人类行为的模型，我们将在第6章就这个话题进行讨论。这个标准模型，也就是通常所谓的完全理性，建立于两个基础的假设之上。第一个假设是人们在经济事务中的主要目的是追求自我利益。经济学家承认，在现实世界里，人们有时也会作出真正利他的行为（尽管这非常难以界定），但他们认为，在简化版中，假设人们普遍会作出符合自身经济利益的事情是合理的。第二个假设是人们在追求自身利益时会采取复杂到令人难以置信且审慎的方式。经济学家常常假设我们在作日常决策时会考虑通胀率、政府未来的预算开支以及贸易逆差等因素。经济学家还假设我们会使用连他们都觉得难解的方程式和运算来处理这些信息。

此外，为了能够预测人类行为，传统经济学家还假设这些超级人类机器人生活在一个理论世界里，那个世界要比人们实际居住的现实世界简单得多。比如，在余生中，当你要作出是否储蓄或是否购买半打啤酒的决策时，如果要将预计利率列入考虑范围，你需要获得关于这些利率的信息。传统模型往往假设，做决策时所需的所有信息可以完整、迅速

地免费获得。当然，实际情况是，我们通常需要根据不完整或不确定的信息作出决策，如果我们需要更多信息，就得付出时间和金钱。其他关于我们生活的世界的典型假设包括：

- 不存在交易费用（比如在买卖过程中没有费用、税、法律约束或其他成本）；
- 所有产品是根据价格进行销售的纯商品（比如没有品牌、产品质量的差别）；
- 公司都在以尽可能高效的方式运营；
- 消费者可以为任何可能存在的不测事件购买保险；
- 经济决策者之间只通过价格进行互动，而且往往是通过拍卖机制互动的。（你常去的超市上一次举行拍卖会是什么时候？）

这些假设加在一起，导致加州大学洛杉矶分校的宏观经济学家阿克塞尔·莱乔霍夫德（Axel　Leijonhufvud）称传统经济学模式是“聪明到不可思议的人处于简单到不可思议的环境中”，而现实世界是“简单到合乎情理的人应对复杂到不可思议的环境”。[17]有无数的证据支撑着莱乔霍夫德的观点。包括赫伯特·西蒙、丹尼尔·豪斯曼和阿莫斯·特沃斯基（Amos　Tversky）[(5)]在内的行为经济学家已经证明，尽管人们在做决策时十分聪明，但这种方式与传统经济学所刻画的大相径庭。[18]现实生活中，人们其实对复杂逻辑计算十分不在行，但是非常擅长快速认知、解读模糊信息以及学习。真实的人还很容易犯错误，并在做决策时陷入偏执。最终，他们会达到赫伯特·西蒙所谓的满意状态，在此状态中，他们寻求的是“足够好了”，而非绝对最好。举例来说，传统经济学会假设在你需要为汽车加油的时候，你会开车到附近的各个加油站，对比找出最便宜的价格。相反，西蒙认为人们会有一个关于价格的模糊概念，他们会选择最近的、看起来价格合理的加油站。[19]我们有理由相信，在一个信息是有成本的、不完整的、急速变化的世界里，人们的大脑会在“足够好了”，而非绝对最优的时候作出快速的决策。

近年来，主流经济学家开始接受传统经济学中的这些不确定性。2001年的诺贝尔经济学奖被授予了乔治·阿克洛夫（George　Akerlof）、迈克尔·斯宾塞（Michael　Spence）以及约瑟夫·斯蒂格利茨，他们的模型指出，并非所有人都能获得完美的信息。2002年的诺贝尔经济学奖被授

予了丹尼尔·卡尼曼（Daniel Kahneman）和弗农·史密斯（Vernon Smith），以表彰他们在更加现实的行为理论方面所取得的成就。包括弗兰克·哈恩（Frank Hahn）与根岸隆（Takashi Negishi）在内的研究者们在“非瓦尔拉斯市场”（例如没有拍卖师的市场）方面做了许多工作。尽管取得了这些进步，我们仍然很难找到同时包含所有这些方面，从而描绘现实环境中的现实人类的模型。[20]均衡是一位严格的掌控者，尽管经济学家有时候会放松一两个假设，但均衡数学的限制意味着真正现实的模型需要从根本上打破传统的框架。

时间不等人

传统经济学因为依赖均衡而付出的代价还包括他们对时间的奇怪看法。大部分传统经济学模型实际上并没有考虑时间因素，它们只是假设经济会瞬间从一种均衡跳到另外一种均衡，而均衡与均衡之间的瞬变条件是无关紧要的。如果模型有时间概念，通常包括“短期”和“长期”两种，或者是一种想象的虚拟索引时间（比如博弈论模型中的回合次数，许多宏观经济模型中的世代数）。很少会有模型可以容纳正常时间概念里的分钟、小时、天和周。[21]然而，时间在现实世界的经济现象中毫无疑问是非常重要的。人们需要时间才能设计、生产、运输、销售产品，获取信息及作出决策。这些事情所需花费的时间将影响我们对于经济动态的理解。

我们可以用一个老笑话来解释这个问题。一个老年经济学家和一个年轻经济学家在街上走路，年轻的经济学家往地下一看，看到一张20美元的钞票，于是说：“看，这儿有一张20美元的钞票！”而老年经济学家回答：“瞎说，如果街上真有一张20美元的钞票，早就被人捡走了。”

以传统经济学的观点来看，如果街上出现了一张20美元的钞票，世界瞬间就会失衡。理性的、利己的人们有动机捡走这20美元，所以会有人走过来把它拿走，使世界恢复到均衡状态。重要的是，我们知道什么才是均衡状态——根本不存在20美元放在街上的情况，人们需要花多长时间找到和捡走钞票以及世界从一种均衡状态转移到另一种均衡状态的具体路径之间并没有实际的联系。

当然，在现实世界里，有20美元掉在路边和有人发现并捡走它之间有一个时间差。这就说明，在任何一个时间点上，路边的某处存在未被发现的20美元是有可能的。在此过程中，时间尺度非常重要，因为掉在

街上的钱数可以成为钱掉在地上与被发现的平均时间的比率的一个函数。通过改变时间尺度，我们就能绘制街上有钱掉落的场景（掉钱速度很快，等待被发现的时间很长）或很少见到的20美元掉在地上的场景（掉钱速度很慢，等待被发现的时间很短），抑或是处于两者之间的情况。我们甚至还可以构建世界在堆满钱和没有钱之间转换的场景。[22]重点在于，除非我们知道相关的相对时间尺度，否则就无法解释系统将如何运作。

缺乏明确的时间尺度是阿尔弗雷德·马歇尔对于100年前的经济学的最大抱怨。在这期间，人们曾作过一些试图将动态引入传统理论的重要尝试，其中包括南加州大学理查德·戴（Richard Day）所做的工作，以及以时间滞后为特征的宏观经济学模型。[23]但由于存在对行为的种种假设，人们几乎不可能创造出能够将均衡和复杂动态以及现实时间尺度结合在一起的模型。[24]

制造有趣的外生因素

如果传统经济学模型没有将时间的概念考虑在内，那么我们就有理由提出这样的问题：它们如何处理经济中出现的变化？之前我们说过，传统经济学理论就像碗中球，无论我们让球滚到碗中的什么地方，它都会回到同一个均衡点上。但是我们知道，经济是高度动态的现象，随着生产能力、价格波动、消费者喜好和技术的变化，各个方面都会随时发生变化。那么，如何才能将这种动态性与均衡内在的静止结合起来呢？如何才能让碗中球随着时间的推移而不断移动呢？

如果我们将碗中球取出，用力震动它，然后将碗的一边压弯，使得均衡被打破，这会出现什么情况？震动一开始就会让球离开均衡点，不断地在碗内滚动。当我们压弯碗的一边（请想象这是一只橡胶碗），改变约束的形状，最终球会根据碗的新形状而在一个新的均衡点停下（见图3-1）。

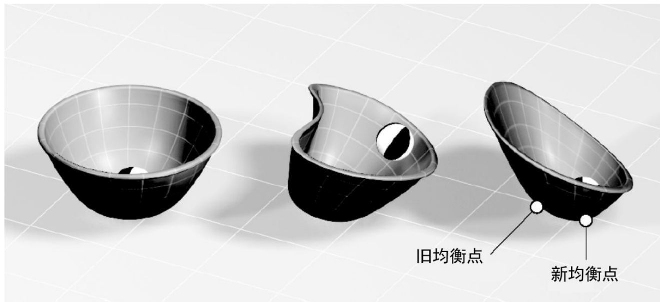


图3-1　均衡、震荡及新均衡

实际上，这正是经济学家在引入外生震动时对模型所做的工作。所有的模型都有限制条件。如果有人想要在模型中加入太多限制条件，那么模型就会变得很大、很复杂，从而失去用处。打个比方，人口增长很明显会对经济系统造成影响，但将生育和死亡因素纳入经济系统却是行不通的。相反，为了简化，我们可以将预估的人口数据作为外部因素输入。如前所述，模型限制之外的变量被称为外生变量，而模型限制内的变量则被称为内生变量。典型的外生变量包括客户喜好、技术创新、政府调控和天气变化。举个例子来说，技术变革（如互联网的发明）可被视为经济系统的外生震动，这一变化会影响生产者的成本（例如戴尔公司可以通过互联网销售而降低电脑的售价）和消费者的偏好（例如消费者更喜欢从亚马逊网站购书）。这些变化又会影响系统的限制条件（碗的形状），从而改变均衡点的位置。由于经济受到外生变量的反复冲击，均衡点也在不断变化。如此一来，经济的动力就来源于均衡、冲击、新的均衡、新的冲击……再次达到新均衡的过程，整个经济会经历一个又一个的临时均衡过程。

这个方法的问题在于，它为经济学家留下了一个摆脱困境的出口，致使他们可以让最困难、通常也是最有趣的问题脱离经济学的约束。比如，如果技术变革被认为是一种随机的外在力量（就好比天气），那么人们就不需要基本的理论来解释技术变革与经济变化之间的关系。[25]同样，人们还可以将经济周期的波动归因于神秘的外在力量，例如消费者的信心或因为新闻而造成的股市崩盘。生物学上也有一个类似的方法。多年来，进化理论家一直在思考物种灭绝的谜团。人类的自然本能是要寻找一个最接近的、相称的原因，也就是说，重大事件必须有重大原因。20世纪80年代，地质学家沃尔特·阿尔瓦雷茨（Walter Alvarez）和他的父亲——诺贝尔奖获得者、物理学家路易斯·阿尔瓦雷茨（Luis Alvarez）提出了一个理论，认为恐龙的灭绝是由白垩纪时期一颗巨大的

小行星撞向地球导致的。实际上，一些证据也支持这个假设。后来，一些研究者选择抽身而出，研究长期的化石记录，结果发现行星理论或许能够解释白垩纪晚期的物种灭绝现象，却不能解释化石记录中其他10种重大的突发灭绝（其中一些事件的规模更大）。后来的研究表明，突发灭绝或许是由进化本身的内在动力导致的，而不是重大的外部事件。[26] 我们将在第8章看到，在复杂适应系统中，微小的内在事件可以在不经意间导致巨大的变化。

在经济学中，外生的“行星”有时确实会撞击整个系统，例如全世界的经济都受到了“9·11”恐怖袭击事件的影响。1987年10月19日的股灾——股市下跌20%又是为什么呢？《纽约时报》当天的新闻标题是“对美元贬值和贸易逆差的担忧”，当然，这样的标题可以用在许多其他重大事故并未发生的日子上。[27]那1982年的经济衰退——12个月内美国公民失业率由7.5%猛升至11%，又该如何解释呢？在那一年里，20位主要的经济学家都预测失业率将会下降。如果重大的外生事件能够引发持续的衰退，为何预测者们预测不到失业率将会上升？[28]在各个案例中，似乎是外生因素在驱动真正有趣的经济行为——某些尚未被完全认识的内在动力正在引发股灾和经济衰退。[29]

人们必须明确模型的界限，但对于一门进步的科学而言，需要随着时间的推移而扩展自身的解释范围。在传统经济学中，均衡的“塑身衣”导致传统模型不得不将某些经济学领域最有趣、最根本的问题排除在外生的围墙之外。

对于事情的限制

许多人都经历过这样的尴尬场景：当你走到麦克风前开始说话，音箱中却传来一阵刺耳的声音——这是正反馈。如果麦克风离扩音器太近，声音在麦克风和扩音器之间循环放大，发出巨大的声响，就会出现正反馈现象。正反馈是一个加速、放大、自我强化的循环。负反馈则相反，它是降速、抑制、自我约束的循环。负反馈的典型案例就是恒温器，如果你的房间太冷，恒温器就会启动加热；一旦温度超过设定值，恒温器就会关闭，直到温度低于设定值。恒温器会抑制房间里的温度波动，使得温度稳定在设定值。

传统经济学假设经济过程是由抑制性的负反馈主导的。我们之前讨论过，这是生产和消费的递减或收益递减。如前所述，传统经济学假

设，生产线上第20位工人的生产力大大不如第10位工人，第5个甜甜圈的吸引力大大不如第1个。跟碗壁一样，负反馈会让事物处于被控制的状态，让事物趋于均衡，并且让世界不被无限多的车辆和甜甜圈淹没。

真实世界确实展现了收益递减的现象。但是，原斯坦福大学和圣塔菲研究所的经济学家布莱恩·阿瑟认为，真实世界同样会展现出正反馈或收益递增的现象。[30]越多年轻人穿运动鞋，运动鞋就越受欢迎。网络上可获取的信息越多，网络本身就越有用。有时候，买某一只股票的人越多，就有越多的人争相追涨。所有这些收益递增现象最终都会逐渐消失。今日的热门趋势就是明日的没落潮流，网络信息会泛滥，股票市场的泡沫最终也会破裂。但我想再重申一次，时机意味着一切。传统经济学理论往往以长期作为假设前提，在长期内，所有的收益递增都会耗尽自己，经济安全达到均衡。但如果没有长期呢？如果一种潮流在另一种潮流尚未开始之前就已经消退了呢？如果有人发明了谷歌帮助我们浏览网页呢？如果有一部分投资人在看到泡沫破裂后依然相信他们能够打败市场呢？在类似于经济这样的系统中，总有新的正反馈资源使气氛活跃起来。现实世界中没有长期，或正如约翰·梅纳德·凯恩斯的名言："长期是对当前事务的错误引导。从长远来说，我们都会死去。如果恰逢狂风骤雨的季节，经济学家只能告诉我们，风暴过后大海将会复归于平静，这样他们给自己安排的任务就显得太容易、太无用了。"[31]

19世纪英国讽刺作家托马斯·洛夫·皮科克（Thomas Love Peacock）在其小说《科罗切特岛》（*Crotchet Castle*）中描述了一场两位绅士之间的辩论，他们争论的是自古希腊以来世界是进步还是倒退的问题。其中一位绅士将经济学的发明当作进步的证据，"科学中的科学"，他如此评论道。反对者则认为，经济学是"一种极其残暴的技术，没有任何一个雅典人能够忍受。假设的前提是要么没有证据，要么完全忽视证据；如果人们从中得出的结论非常合乎逻辑，那么这些结论必然是错误的"。[32]

为了让复杂、动态的世界经济能够被装进简单、静态的盒子里，经济学家不得不采用"没有证据的假设"，这使得模型基于假设得出的结果出现了严重问题。如果没有这些假设，简化单一的碗中球模型就会失去作用；碗的平滑底部能够产生碰撞，球永远不会停下，方程不能被解决，均衡的可预测性就会消失。传统经济学家或许会找到这样的理由来回复："这些年来，我们为建立这些模型付出了大量的努力，我们已经通

过放松部分假设而加强了其现实性。”经济学家利用不尽完美的理性、不完美的信息、市场摩擦和动态，利用内生办法处理原本外生的变量，从而建立了模型。有人会问，传统经济学的模型是在何处同时放松了所有这些限制，因而看起来像一个动态系统的？为了回答这个问题，提问者必须放弃经济是一个均衡系统的观念，而传统经济学直到近来也没打算这么做。

现实验证

假设我们相信米尔顿·弗里德曼，无须对传统经济学的假设“一探究竟”，那么传统经济学在弗里德曼的实证检验中表现如何呢？

许多非经济学家对经济的科学凭据持怀疑态度，这是因为经济的预测结果不尽如人意，比如对经济增长、利率和通货膨胀的预测。然而，我们应该记住，科学的标志不在于预测未来的能力，而在于解释问题的能力，它能增强我们对于宇宙运转机制的理解。[33]如前所述，科学在预测方面所起的作用是帮助我们在不同的解释之间作出分辨。一个成熟的理论往往拥有经得起测试的逻辑启示。比如，气象学家和气候学家能够告诉我们许多关于大气化学过程的知识，他们的理论可以通过在北极钻取冰岩芯和向高层大气发送气象气球来验证。然而，这并不意味着他们就能十分肯定地告诉你，下周日你去郊游的时候是否会下雨。这是因为地球的大气跟经济一样，是一个复杂而又高度动态的系统，距离均衡状态非常遥远。事实上，气象学和气候科学能够解释为何天气预报天然就是不准确的。在科学的各个领域，研究者都能解释现象并且验证解释的正确性，但未必能作出准确的预测。举个例子来说，生物学家能够解释却不能预测蛋白质的折叠，物理学家能够解释却不能预测湍流的确切运动方向。

科学是一个持续的学习过程，在此过程中，科学家会对相互矛盾的观点进行验证并在此过程中积累知识。正如卡尔·波普尔爵士（Karl Popper）在20世纪30年代指出的那样，没有“终极证据”证明某个理论是正确的，但是我们可以说某个理论是否被数据证明是不成立的，某个理论是否比其他理论与数据更吻合，某个理论目前还与数据不吻合。[34]比如，我们不能说爱因斯坦的相对论已经被证明了，但是可以说它的预测已经得到了充分的验证，目前尚未出现矛盾之处，并且相对于其他解释，它与数据的吻合度更高。科学理论必须经历一个提出各种解释、严

格地用可以被验证的方式将它们表达出来，去除那些无法经受验证的理论并把成功经受验证的理论作为基础的过程。[35]

在此标准之上，我们就能提出一个问题：传统经济学的预测与数据的吻合度有多高？答案是“并不是很高”。法国高等社会科学研究院教学主管、备受尊敬的微观经济学家艾伦·科曼（Alan Kirman）评论道：

> 过去100多年间创建的大多数优雅的经济学理论框架将在未来10年被验证——在均衡是由什么构成的这一理念上，它们提供了错误的聚焦、误导以及昙花一现的观念。如果我们认为科学理论应符合两种标准规范——预测和解释，事实证明，经济学理论至少是不充分的。现实生活中，几乎没有人对经济学理论的差劲预测能力表示满意。许多人为此进行了辩护，但结论甚至不值得提及。许多经济学家认为，我们对于经济学运作机制的理解还在提升，因而可以质疑经济学理论在这方面并不充分的说法。然而，证据并不能打消人们的疑虑。人们对于包含经验数据的经济学理论几近病态的反感已经到了这样的程度——“解释”被当作一种独立的而非可测试的概念。[36]

正如科曼所言，除了数据与传统经济学理论相悖的问题，许多理论甚至根本没有经过验证。经济学的一个分支，所谓的计量经济学是负责数据分析的。[37]大部分计量经济学的工作不是验证理论模型，而是致力于发现变量之间的统计关系（通常是为公共政策或其他应用目标服务的）。事实上，统计关系并没有为现象提供因果解释。正如许多经济学家所指出的那样，我们常常缺乏可用的数据来验证理论。即便有可用数据，通常也是噪音数据或其他问题数据（尽管经济学家从物理学家或生物学家那儿没有得到什么同情，但这两类科学家不得不为了这些理论而建造粒子对撞机和太空望远镜，以及绘制人类的基因图谱）。

然而，传统经济学在两个方面经受住了严格的验证。首先是金融理论，金融市场的实时数据和巨大的计算能力为传统经济学理论的验证提供了前所未有的条件。不幸的是，对于传统经济学来说，这些数据不断被证明与理论得出的最基本的预测相悖，我们将在第17章对此话题进行讨论。[38]其次是实验经济学。有一种盛行的观点认为，经济学与其他科学的不同之处在于人们无法通过对照实验来验证假设。比如，美联储不能无限提高利率来检验到何种程度时会引发经济衰退。然而，尽管针对整体经济进行实验存在很大的困难，但我们却可以对经济的微缩版进行

实验。研究者让不同小组的人互相议价、在拍卖中竞标、玩经济游戏、对模拟股票市场进行调查、去实验商店购物，并且参与建构各种各样的人为环境，以便捕捉人类经济行为的各个具体方面。这样就产生了丰富的成果，我们将在第6章对其中的某些方面进行讨论。需要再次申明的是，通常来说，传统经济学的许多核心观点都很难获得数据。[39]

当然，大部分关于传统经济学的验证工作都具有很高的技术性要求（对此感兴趣的读者可以参考注释）。尽管如此，在接下来的几个部分，我们还会看到多个直观的案例，看到传统经济学的关键预测并不能满足科学现实验证的标准。

供求“定律”

如前所述，传统经济学最古老的原则之一就是供求定律。关于这个定律的基本预测是，供应和需求的平衡力量会驱使市场达到一个平衡的价格和数量水平，实现市场出清。这个理论与现实情况相当接近，效果很好。举例来说，如果一个汽车公司推出了一个新的车型并很快受到热捧，公司通常会在需求大于供应时提高价格，然后扩大生产；一旦需求下降、供应增强，他们就会降低价格。一切都如理论所预测的那样。

但如果我们进入一个更加细颗粒度的层面，就会看到现实世界的市场几乎永远不会达到均衡状态，供应很少与需求相等，市场很难实现均衡。几乎所有市场都是建立在非均衡而不是均衡的假设基础上的。大多数市场都有库存、订单储备、过剩的生产能力以及起缓解不均衡状态作用的中间商。一个地方的汽车经销商有一停车场的车卖得很慢，也有许多“热门”车型的订单，消费者都翘首以盼。一个地方的超市也几乎很少达到均衡状态，它的库存时高时低，里面充满着从后门运进来的食物供应和从前门运出的需求之间的不平衡。服务行业，例如律师和会计，很少能够百分之百地发挥他们的专业技能，因而其中都为需求的波动保留了一些可用的摇摆产能。甚至于金融市场——经常被看作最接近理论典范的市场，也拥有应对不可避免的供求不平衡的机制，在纽约证券交易所有“特种经纪人”，在纳斯达克交易所有“做市商”，他们的工作都是为了缓解供求之间的不平衡。

供求定律其实并不能称之为定律（至少从科学的层面来说不能），它更适宜被称为“关于供应和需求的粗略估算”。部分传统经济学家或许会分辩说，库存和产能过剩并不是问题，它们只不过是供应-需求平衡中

的杂音。事实上这种说法是错误的。在本书的后面章节我们将讨论到，库存和产能的动态能够帮助我们解释一些现象，例如价格波动和商业循环现象，传统经济学通常利用外生力量来进行解释。[40]正如我们将在第8章看到的那样，股票市场中的库存甚至有助于解释股票市场的波动。[41]

一价定律

在传统经济学中，第二条有名的“定律”是一价定律，其内容是：“在没有运输成本和贸易壁垒的情况下，同一商品在所有市场中的售价必须相同。”[42]例如，纽约的金价应该与伦敦的金价相同，如有差异，应将其归因于从一个市场到其他市场之间的运输成本以及进口关税、税费及其他贸易成本，否则就有人能从差价中套利，即通过在低价市场买入、到高价市场售出，获得无风险利润。套利者的买入和售出会导致两个市场的价格复归均衡。跟供应和需求一样，一价定律通常也不过是一种近似的状态。高流动性的商品（例如金子）通常在不同的市场中很少会出现“可套利”的情况。

即便如此，这种近似作用通常也会发生故障，无论是在整体的宏观经济层面还是更加微观的个人生产和服务层面，都是如此。对此理论的一个重要验证就是大部分欧洲经济体加入欧盟，尤其是引入欧元的案例。理论预计，大规模消除贸易壁垒、增加流动性、降低货币交易成本、增强价格透明度的行为，必将导致欧盟范围内价格的统一性增强。但事实证明，结果恰恰相反。欧盟统计局的数据显示，自1999年引进欧元以来，各国间的价格差异在不断扩大。欧元区的价格标准偏差从1998年的12.3%扩大到了2008年的13.8%，这与理论预测的正好相反。欧盟内部市场委员会经济分析主任弗朗西斯科·卡巴莱罗-桑斯（Francisco Caballero-Sanz）认为，价格不统一的原因在于消费者并不像经济学理论所“希望”的那样理智。[43]

从更细致的层面来看，我们通常看到个人商品和服务在价格上存在很大的差别。比如，来自德累斯顿嘉华投资银行的全球股票策略总监詹姆斯·蒙蒂尔（James Montier）对伦敦番茄酱市场进行了一番异想天开的调查，然后发现同样的番茄酱在区域内各超市的价格差异明显，[44]实际价格与理论预测的价格差别最高可达43%。在这种情况下，伦敦番茄酱市场就存在很多无风险的获利机会。在传统经济学的理论世界里，这样的机会会因为套利而立即消失：但在现实世界里，人们需要时间来发

现这些套利机会，机会时隐时现，有些值得利用、有些不值得，在利用的时候或许又存在各种各样的障碍。

一些传统经济学家或许会问，如果一价定律只是一种理想状态，这又有什么关系呢？然而，对于大多数科学家而言，欧元区内存在13.8%的标准偏差，或者番茄酱价格存在43%的价格误差，这种理想化也太过于宽泛了。关键不在于一价定律背后的基本理念是错误的，人们在“无壁垒的情况下”当然有意愿从价格差中套利；而是在现实世界里，各种各样的壁垒几乎永远存在，无论是基于以下哪种事实：有的人没有时间寻遍所有商店从而找到区域内价格最低的番茄酱，或者各种各样的交易成本以及运输、法律和其他影响欧盟内交易的问题依然存在。事实上，科学家们所感兴趣的关于价格收敛的问题，是套利动机与各种壁垒的无常变化之间长期的、动态的相互作用的结果。然而，均衡框架在数学方面的要求迫使经济学家剥离了这种复杂性，只留下一个预测结果的价值尚且存疑的“定律”。更加有用的定律应该是能够应对现实世界中的价格复杂性的理论。[45]

数百万兆年里的均衡

传统经济学最基本的预测或许是：经济作为一个整体必将在某个点上达到均衡状态，这是微观经济学的一般均衡理论和标准宏观经济学都会作出的预测。[46]如前所述，传统经济学并不认为均衡是一种永恒的状态，而认为经济将经历一连串的冲击、均衡、冲击、均衡的动态过程。我们可以想象这就好比有人猛烈冲击碗中球，看着它静止下来，又再次猛烈冲击它，如此反复。但是对于想要达到均衡状态的系统来说，两次冲击之间的时间必须长到让球能够静止下来。若非如此，球就只会永远不停地滚来滚去，然后又被随机地连续撞击，永远无法达到均衡状态。

如此一来，关键问题就在于，经济达到均衡状态需要多长时间？何时才是它的“静止”时间？20世纪70年代，耶鲁大学的经济学家赫伯特·斯卡夫（Herbert　Scarf）指出，达到均衡的时间根据经济中产品和服务的数量呈四次方的指数级增长。[47]这种关系背后的直觉是直接的：产品和服务的数量越多，所有市场之间进行互动所需的时间越长，所有价格和数量进行调整所需的时间也就越长。如果按照第1章所说的那样，粗略估计现代社会大约有10^{10}个库存单位，我们乐观估计，经济领域的每一个决策都以超级计算机（当前是IBM的蓝色基因，每秒可进行70.72万亿

次浮点运算）的速度进行，按照斯卡夫的理论，经济每经历一次外生冲击，需要4.5百万兆年（4.5×10^{18}）才能恢复到一般均衡状态。鉴于每一秒都有各种各样的因素（例如技术变革、政治变革、气候、消费者喜好变化）冲击着经济，而宇宙现在仅有120亿岁（1.2×10^{10}），这确实是个问题。[48]

非随机游走

如前所述，传统经济学中最有可能经受得住现实验证的或许要数金融理论。但是，尽管传统金融理论的预测不像第一近似值那么糟糕，但一旦经受近距离检视，它们也会崩溃。

传统金融的一个最有名的预测是股票的价格会遵循一种“随机游走”的规则。我们将在后续章节进行详细讲解，但总的来说，随机游走意味着价格的变动没有模式可循，回顾过去的价格并不能为未来的价格走向提供线索。乍看之下，股票价格确实很像随机游走，尤其在相对平静、表现“正常”的情况下。数十年来，研究者们始终相信，价格的确是随机的。近些年来，通过更优的数据和更强大的工具，人们分析发现价格并不遵循随机游走规则。比如，来自麻省理工学院的罗闻全（Andrew Lo）与来自宾夕法尼亚大学沃顿商学院的克雷格·麦金利（Craig MacKinlay）用1962年至1985年间1 216周的股票价格数据验证了随机游走假设。他们验证了单支股票、股票组合和股票指数，所有这些案例都与随机游走假设相悖。[49]许多使用其他样本和其他方法的研究结果也与随机游走假设不符。[50]

有趣的是，在市场出现重大波动时（也就是说距离均衡状态最远的时候），股票与随机游走之间的不吻合在数据上表现得最为明显。[51]股票价格数据里也有明显的动态结构和信息，尽管关于是否有人可以有组织地根据这些信息赚钱尚存争议，但随机游走的预测是错误的这个事实对于增强金融科学的可信度而言毫无用处。[52]

传统经济学理论的预测也不完全是疯狂的。供应确实近似地等于需求，价格有时（并不总是）确实会收敛。市场或许永远不会真正地达到均衡状态，但是可以以貌似均衡的形式运转。从表面上看，在金融市场

稳定、运行良好时，它似乎确实遵循着随机游走，也就是说，稳定和运行良好的状态被打破时则不然。如果你觉得这些陈述听起来并不科学，那是因为它们真的不科学。正如弗里德曼的对手赫伯特·西蒙所言：“可以肯定的是，经济学已经发展出了大量高度复杂的数学定律，但其中大多数定律与经验现象相距甚远，并且通常只能显示被观察对象之间的定性关系。”[53]

传统经济学建立在摇摇欲坠的假设的基础之上，这就导致了它的结论也站不住脚。下一个问题是，为何传统经济学沦落到了这个地步？传统经济学的问题出在100多年前，出在瓦尔拉斯的一个关键举动上——将均衡的概念从物理学引入经济学。瓦尔拉斯并没有意识到自己从根本上对经济学进行了错误的分类。

被误用的比喻

人类能够熟练地识别模型，使用比喻来加强对于世界的理解和推理。指出一种事物与另一种事物的相似之处或相同之处，能够帮助我们快速地抓住一个复杂现象的本质。莎士比亚本可以用长篇巨著向我们展示朱丽叶对于罗密欧的生命来说有多么重要，如何给他带来了幸福，等等。但事实上，莎士比亚只用了一个简短的句子“朱丽叶就是太阳！”表达了这些含义，并且含义更加丰富、形式更为有力。

科学家同样也会使用比喻，以便激发创造力，促进关于复杂理论的沟通。举个例子来说，弦的小型震动回路的比喻启发物理学家提出了弦理论（此理论试图统一宇宙的根本力量并解释亚原子粒子的来源），使得他们发现了与前辈完全不同的思考方式。[54]同样，相对于“十一维的卡比拉-丘空间”，弦回路一词使得物理学家能够以比喻的方式，更容易地与外行听众沟通弦理论的核心概念。但是，尽管比喻有助于科学交流，但科学本身不仅仅是以比喻为基础的。科学理论的内容不仅仅包含某样事物与其他事物相似。正如我们在第1章中讨论过的那样，科学家会提出某事物确实属于某种普遍现象，但相似不等于相同。宇宙学家说太阳是恒星，这不仅仅是说它在某些方面与恒星相似，而是说太阳是一种普遍现象，即所谓的恒星中的一员，它拥有某些在现实中可观测到的特征。[55]

毫无疑问，瓦尔拉斯在阅读潘索的物理教科书时，受到了物理学系统中平衡力与经济学系统中平衡力概念之间的相似之处的启发。[56]这种

相似激励了他将均衡分析的数学工具运用到经济系统中。100年后，在圣塔菲研究所的科学家与经济学家会议上，摆在台面的关键问题就是，经济学中均衡概念的提出仅仅是基于物理系统与经济系统之间的表面相似性，还是经济系统确实就是均衡系统，因而拥有系统的普遍性质？[57]换句话说，传统经济学中的均衡框架是一种比喻还是一门科学？

对于瓦尔拉斯及其支持边际主义的同事来说，这样的问题本不应该被摆在他们面前。首先，当时人们对于科学哲学以及对于比喻的合法性和非合法性的理解尚未到达如今的程度。其次，这个问题还会引出另一个问题：如果经济系统不是均衡系统，那又是什么系统呢？第二个问题同样也会让瓦尔拉斯感到困惑。对于瓦尔拉斯那个时代的物理学家来说，系统只有两类：均衡系统——可以用数学方法进行分析的系统；以及“其他”系统——其中部分可以用数学方法分析，但大多数都不可以。对于类似经济这样的现象，当时没有很多别的办法对此进行分类。由于瓦尔拉斯的目标是将数学的严谨与可预测性带入经济学，所以，他最终踏上了均衡分析这条多人踏足的道路也就不足为奇了。

尚未成熟的物理学

瓦尔拉斯在履行将经济学转变为一门科学的使命时，科学本身尚处于发展阶段，缺乏一些关键概念。所有科学都在不断进步，边际主义者所了解的物理学并不是我们今天能在教科书上看到的经典热力学；事实上，我们可以说它在当时尚未成熟。边际主义者从物理学中借用的内容包括热力学第一定律，但是他们忽略了第二定律。

热力学第一定律认为，能量既不能被创造也不能被消灭，因此也被称为能量守恒定律，它发展于19世纪早期到中期，瓦尔拉斯、杰文斯和其他人在读到相关材料时这个定律已经非常清晰。有关第一定律的作用，我们可以想象有一个小球被拿到了碗的上方。当球掉落，它所拥有的势能就会被释放出来。如果你把手放在碗的底部，就会感受到球由于重力而对碗底造成的冲击。能量的释放被称为功（因为它有可能有用武之地，例如驱动机器）。球坠落之后会开始滚动，它会受到碗的摩擦力，而摩擦会产生热量。因此，球的势能就会被转换成功和热。伟大的英国物理学家詹姆斯·普雷斯科特·焦耳（James Prescott Joule）指出，自然界对于能量极其吝啬，能量既不能被创造也不能被毁灭，仅能从一种形式转化为另一种形式。如果将在碗的最高点时储存于球内的势能与

滚落时所做的功及释放的热量相比，这两者的数量应该是相等的。同样，如果测量存储于一团煤炭中的能量及燃烧煤炭所做的功（如驱动火车头），存储于煤炭中的能量应该等于它所做的功以及从火车头烟囱中排出的、被浪费掉的热量。[58]

第一定律的属性之一在于，如果某个系统的总能量是固定的，也就是说“守恒的”，那么系统最终必然会趋于均衡。一旦球的势能不再转变为功和热，那么球就会在碗底以最低的能量状态静止。同样，一旦燃烧的煤炭被转化为了功和浪费的热量，它就不再燃烧，达到均衡状态。只有加入外部的能量，例如晃动碗或加入更多的煤炭，我们才能破坏此时系统的均衡状态。

正如巴黎圣母院的菲利普·米罗斯基（Philip Mirowski）指出的那样，瓦尔拉斯借用均衡概念的后果之一，在于引发了传统经济学模型对于固定或守恒数量的数学需求。因此，传统经济学通常将价值描绘为能从一件物品转移到另一件物品的固定数量（例如，从资源转移到商品可以换来钱；钱换回商品，然后进行消耗，创造实用性）。[59]实际生活中并没有新的财富被创造出来，世界上有限的资源先是被分配给了生产者，他们又创造出有限的商品，这些商品被分配给消费者。人们可以用或高效或低效的方式来分配财富，就像他们可以用高效或低效的方式烧煤，但在一般的均衡模式中，经济无法创造出新的财富，就像煤不能被循环利用。[60]由于人们强调财富蛋糕的大小是固定的，英国经济学家莱昂内尔·罗宾斯（Lionel Robbins）在1935年提出了将经济称为“稀缺的科学”这一著名的说法。[61]这在现代经济学教科书中仍有体现，比如，在保罗·萨缪尔森和威廉·诺德豪斯（William Nordhaus）的书中，经济学就被定义为“关于社会如何利用稀缺资源来生产具有价值的商品以及如何向人们销售商品的研究”。[62]曾经给瓦尔拉斯和杰文斯带来启发的第一定律的遗留问题至今留存于传统经济学之中。

然而，第一定律仅仅是故事的一半。第二定律没有出现在瓦尔拉斯和杰文斯当时所了解的物理学中，第二定律指出了衡量系统混乱或随机程度的指标——永远都在增长的熵。宇宙作为一个整体不可避免地会从有序状态转变为无序状态，其最终的归宿是一种随机的、毫无特征可言的、温度完全均匀的瘴气。随着时间的推移，宇宙中所有的秩序、结构和模式都会瓦解、衰退、消失，如汽车生锈、建筑崩塌、山峰被侵蚀、苹果腐烂、倒入咖啡里的奶油消散至完全融合。熵为时间指明了方向。伟大的物理学家、圣塔菲研究所的创始人之一默里·盖尔-曼指出，如果

你有尚未长大的孩子或孙辈，你的食品柜里可能既有花生酱又有果冻。随着时间的推移，在给孩子们做三明治的时候，花生酱会混入果冻中，果冻会混入花生酱中。盖尔-曼指出，如果他让你看花生酱中如何越来越多地混入果冻的延时视频，然后再让你看一个果冻消失、与此同时花生酱变得干净的视频，你立即就会明白哪一个影片是正向播放，哪一个是倒退播放。[63]如果没有熵和不可避免地从有序到无序的转变，我们就无法分辨什么是过去、现在和未来。从熵的发现开始，它就成了物理学家看待宇宙的一个核心概念。[64]

但对于瓦尔拉斯、杰文斯和其他传统经济学的建设者来说，这条自然界的最高法则并未体现在他们的框架中。第二定律是一个经过了长期酝酿的科学概念，包含了1824年到1865年间许多人的努力，其中包括萨迪·卡诺（Sadi Carnot）、鲁道夫·克劳修斯（Rudolph Clausius）和威廉·汤姆逊爵士。然而，它的重要性直到19世纪末才被人们意识到，而它的许多重要含义直到20世纪才开始发挥作用（一直延续至今）。熵这个概念对当时的人来说太新了，人们对它的理解过于浅薄，以至于未能将它加入当时给了瓦尔拉斯和杰文斯以启发的入门教科书中。[65]

芝麻开门

在对热力学的第一定律和第二定律有所了解之后，我们就能开始接触另一个在瓦尔拉斯和杰文斯的时代无法触及的概念：关于封闭系统和开放系统的概念(6)。热力学系统是指任何一套我们希望界定和研究的确定的空间、物体、能量或信息。宇宙本身就是一个系统，在宇宙这个最大的系统之中，人们可以定义任意数量的小系统。比如，地球是一个系统，你的身体、房子、装满水的浴缸也都是系统。封闭系统是指与其他系统没有任何互动或交流的系统，没有物质或信息进入或流出，宇宙就是一个封闭系统。宇宙没有“外界”，没有其他系统存在于它的边界之外并且可以与它互动。[66]能量可以被转化成物质，反之亦然，而且系统内的能量有可能会被转化成不同的形式，但根据第一定律可知，总量是恒定的。此外，随着有序状态转变为无序状态，封闭系统最终会趋于均衡，封闭系统内的总熵永远都在向着最高水平增长。

第二类系统是开放系统，能量和物质会不断进入、流出。这样的系统能够利用流动的能量和物质来暂时对抗熵，并创建一段时间内的秩序、结构和模式。举个例子来说，地球就是一个开放系统，它处于太阳

流出的能量河流的中央。这种能量流创造出了大型的、复杂的分子，分子又创造了生命，进而创造出一个秩序与复杂性共存的生物圈。熵并没有置身事外，地球上的事物一直在瓦解和衰减，所有有机物最终都会死亡。但来自太阳的能量会不断地为新秩序的诞生提供能量。在开放系统中，由能源赋能的新秩序创建与由熵驱动的旧秩序瓦解之间的斗争从未停歇。自然的会计规则十分严格，在开放系统中，创建新秩序需要付出代价。为了在宇宙某个地方创建新秩序，其他地方的旧秩序就必须被摧毁，这是因为净效应必须永远是熵的增长（秩序的减少）。因此，当太能的能量在地球上创建新秩序时，新秩序里所有的生命和活动都会产生热量，热量又会被反射回太空。无论热量最终去到哪里，它都会产生一种随机效应，从而导致熵增加。地球就是通过这种方式来引入能源并输出熵的。

假设你一直忙于工作，第二定律让你的房子从有序状态变为无序状态。你决定投入一些能量来对抗熵——整理房屋。你燃烧卡路里来擦洗、归置物品，从而向房屋注入能量，同时你还会用电来驱动吸尘器、洗碗机和洗衣机。此外，物质还会以一种高度有序的状态进入房屋，包括食品、衣服、清洁用品等。然而，当你和你所使用的所有物品向周围的环境辐射热量时，宇宙就能获得熵。并且，你用的电也会产生浪费的热量，发电站的烟囱排放物和离开你的房屋重新进入环境中的物质都以垃圾的形式处于一种无序状态。你的房子这个系统引入了能量和物质，并在房子的范围内被用来创造秩序，然后又向宇宙的其他部分输出热量和无序物质，从而输出它的熵。

封闭系统永远都有一个可预测的最终状态。尽管它们在进程中或许会做一些无法预测的事情，但最终都会趋于最大值的熵的均衡。开放系统则要复杂得多。有时，它们会处于稳定的、看似均衡的状态，或者它们会展示出非常复杂而不可预测的行为模式——距离均衡状态十分遥远，这些行为模式包括指数增长、彻底崩溃以及大起大落。只要开放系统能够获得自由能，它的最终状态或者说何时能够达到最终状态就无法预测。

在第1章中，我将复杂适应系统定义为互动的行为主体能够相互适应并且适应环境的系统。复杂适应系统是开放系统的一个子类，它通过能量来处理信息、维持秩序、创造复杂模式。比如，蚁群通过带回蚁巢的食物和物品获得能量和物质，在建筑蚁巢和组织活动的过程中，它们利用能量和物质来对抗熵。自由能量的存在使得复杂适应系统（例如蚁

群）可以远离均衡、创建秩序并且处于动态之中。如果你移除这种能量，那么熵就会占上风，系统就会衰退并最终趋于均衡状态。作为圣塔菲会议的参与者之一，密歇根大学的理论家约翰·霍兰德曾说过：“事实上，系统一旦达到均衡，它并不是稳定了，而是死了。”[67]

经济的错误归类

瓦尔拉斯与杰文斯并不熟悉第二定律，因此他们并没有意识到开放系统与封闭系统之间的差别，也没有意识到复杂适应系统的存在。人们对于开放系统的细致了解在20世纪才真正开始，19世纪60年代至70年代，在生于俄罗斯的化学家伊利亚·普里高津（Ilya　Prigogine）的推动下，相关研究开始加速。当时，传统经济学的模型是基于一个隐含的假设建立起来的，即经济是一个热力学的封闭均衡系统，瓦尔拉斯、杰文斯及其他边际主义者甚至不知道他们将这个假设加入了理论之中。在后续的100年里，经济学和物理学分道而行，这个假设就埋藏在了传统经济学的数学核心里。[68]

在经济学理论方面，对于第二定律的无意识意味着边际主义者及其后辈在根本上（虽然是无意地）将经济进行了错误归类。经济并非封闭的均衡系统，它是一个开放的非均衡系统，更确切地说，是一个复杂适应系统。而证据就在你的窗外，就在你的鼻子下方。相关证据是如此明显，但它之前逃过了大部分经济学家的注意，直到近期才有所改观。[69]如果说经济是封闭的均衡系统的话，那么它的定义特征是趋向于更少的秩序、复杂性和结构，因为熵会让任何封闭的均衡系统不可避免地走向毫无特征的静止。封闭的均衡系统不能即时自我管理，它们不能产生模式、结构和复杂性；最重要的是，它们不会随着时间的推移而创造新的东西。[70]在你的窗外，与经济相关的所有运动、组织和活动不可能是封闭均衡系统的产物。正如我们在第1章看到的，从能人的10^1库存单位到现代社会的10^{10}库存单位，经济的定义特征变成了复杂性的剧增。从石器时代至今，经济活动的增长已经成了一个大张旗鼓地对抗熵的长篇故事，而这种故事仅会发生在开放的非均衡系统中。

传统经济学家或许会问，开放系统和封闭系统的物理性与经济学到底有何关系？开放系统和封闭系统是物理学概念，而经济学是一种社会现象。我是否也像我所批判的瓦尔拉斯那样，使用了不正确的比喻推理？答案是否定的。我们必须记住，社会系统不仅是存在于经济学家脑

中或教科书方程式中的抽象数学模型。它们是由物质、能量和信息组成的真实的物理系统，它们是由人和窗外的一切事物组成的，并且它们跟其他现象一样，也遵循着物理学定律。每天都有大量真实的、物理的能量注入真实的、物理的经济体，这就是经济体运转的动力所在。[71]我们的狩猎采集祖先向他们的经济领域注入的能量是他们所食用的食物和收集的柴火，现代经济体为自身注入能量的方式是巨无霸、微波炉即食餐，还有石油、天然气、煤炭、水力发电、核能源以及其他任何人类可以利用的能源。进入经济的能源为对抗熵和创建秩序提供了能量。同样，经济遵循第二定律，在排放垃圾、污染物、温室气体和热量的同时，它们也在向宇宙输出无序状态。

经济不仅与开放系统极为相似，在事实上、物理上确实也是宇宙开放系统群体的成员。如果我们切断经济的能量供应——断绝食品、石油、天然气和煤炭，那么熵就如入无人之境，经济就会趋于真正的均衡。可悲的是，我们今天仍能看到这样的情况，许多被战争蹂躏的经济体不可避免地会衰退，因为熵会占上风，这些地方的经济会趋于真正的均衡——那是充满痛苦与饥饿的均衡，而从定义上来看，不断发展的、有活力的经济从不会走向均衡。

甚至于只由“纯粹信息”组成的、仅存在于交易者电脑屏幕里的市场（例如纳斯达克股票市场）也遵循着物理世界的某些原则。几年前美国就出现了一位出乎意料的“提醒者”——一只松鼠钻进了康涅狄格州一处临近纳斯达克主机中心的变电器里。这个小东西受了电击之后切断电源，直截了当地关闭了市场。[72]正如我们将在第14章讨论的那样，所有创造财富的经济活动都需要某种能量以及各种各样的物质和（或）信息（尽管反过来说不一定是对的，即并不是消耗能源、利用物质和信息的活动就一定能创造财富）。经济活动严格根植于现实的物理世界，因此经济学理论也不能脱离热力学定律。

传统经济学家或许会承认经济是真实的物理系统，甚至承认它们是开放系统，但他们依然会分辩说开放系统和封闭系统的物理性质无关紧要，因为相对于物理基础来说，经济学家对经济模型进行了高度的抽象化。也就是说，经济模型比物理科学模型更加粗粒化。就像生物学家不参考亚原子物理也能得到细胞模型，经济学家也可以在不明确参考物理热力学的基础上得出经济模型。经济学有着属于自己的概念，例如消费者偏好、价格和产品功能，它们是作为社会领域的抽象概念而存在的，这就好比生物学家有着属于本领域的高度抽象的概念（如基因）。传统

经济学家或许会提出，瓦尔拉斯和杰文斯从物理教科书中借用了一些数学技巧，他们的罪过之处不过就是试图让经济学的抽象概念更加严谨。均衡分析只不过是一种数学技术，仅此而已。那么，既然物理学和经济学都可以使用代数，为何经济学使用均衡的概念就不行呢？[73]

其实，这个赞成均衡的观点中包含着两层意思。第一是“不同程度的抽象化”的观点，第二是“只不过是数学而已”的观点。当然，对于不同的现象，科学需要不同程度的抽象化。如前所述，科学理论既可以是像高速公路地图那样的大图像、粗颗粒，也可以是像区域街道地图那样的细颗粒。它们都同样有效，也必须互相兼容并符合现实。因此，经济学可以有属于自己的高度抽象、粗颗粒、不属于物理学的概念，这些概念可以不对物理定律进行详尽的参考。然而，经济学理论却不能与基本的物理定律相悖。“经济是一个封闭的均衡系统”的说法明显不符合基本的物理学定律。为了使“抽象化的程度”这个论点站得住脚，我们必须相信，即便经济实际上是一个开放的不均衡系统，由于某种原因，它也最好被当作一个封闭的均衡系统。事实上，反过来才是对的。我们将在本书第二部分看到，不现实的假设和缺乏实证的问题是这样产生的：封闭的均衡技术对开放的不均衡经济进行了错误的模型化。[74]

赞成均衡论点的第二个层面——“只不过是数学而已”，错误地理解了均衡是什么以及数学是什么。数学是一种语言，它是一个我们用来描述和解释世界的符号系统。[75]数学是一种特殊的语言，因为人们对于各种数学符号所代表的含义以及使用这些符号的规则达成了高度的一致。人们进行过各种各样的尝试，试图创造一种“纯粹”的数学——它仅仅是一种抽象的语言，其符号仅仅是纯粹的逻辑结构，与现实世界无关。[76]尽管这样的尝试对于拓展数学领域的边界大有裨益，但当前的数学家和哲学家都认为没有人可以把数学与它在物理世界中的意义剥离开来。[77]当然，数学家可以创造一个在物理现实中并不存在的虚幻世界，例如不可见、不可触摸的43面体。但我们用自然语言也能这么做——我们可以通过写小说或制作电影来创造现实世界中并不存在的事物。然而，人们对于这些虚幻事物的理解，无论是最新的好莱坞大片，还是43面的超球面体，都不可避免地会跟他们在现实世界的经验产生关联。[78]并没有什么事情“只与数学有关”。

如果“经济均衡”在现实世界中代表着某种含义，那么接下来的问题就是，它代表的含义是什么？并没有什么规定要求两个不同领域的研究者必须按照相同的含义来使用术语。因此，我们要问“均衡”这个术语在

经济学和物理学中的含义是否相同。经济学家对于均衡的定义或许全然不同，那么我所提出的问题就变成了一个术语混淆的问题。如果均衡在不同领域的含义不同，那么我们应该可以由一种定义翻译出另外一种定义。然而，据我所知，经济学中的均衡的定义与物理学中的均衡的定义没有任何区别。恰恰相反，这两个领域对于该词的运用完全一致。《牛津物理学词典》对于均衡的定义是“系统的（一种）状态，在这种状态中，力、影响和反应等相互平衡，因而不会发生任何净变化”。[79]而《科林斯经济学辞典》将均衡定义为“一种没有变化趋势的平衡状态”。[80]此外，针对均衡分析，经济学家和物理学家所使用的数学方法也是一样的。

这就引出了一些令人不安的问题。传统经济学真的说过经济学和市场是均衡系统吗？这是一种合理的理论延伸吗——我们是否可以将一系列广泛的类别都称为均衡系统，包括吊在弹簧上的砝码以及猪肉市场？然而，经济学又怎能被称为封闭的均衡系统呢，是因为能量和物质从它的一端进入，熵从另一端输出吗？这种说法难道不违背物理学定律吗？还是说，经济和市场仅仅是与均衡系统类似？这会不会是由于特殊的经济学历史而导致的错用比喻呢？[81]

我认为基于传统经济学核心的新古典主义模型是建立在一个错用比喻之上的。19世纪晚期的经济学家并没有意识到这一点，他们好心好意地从物理学借用了一系列理念，从而从根本上将经济错误地归类为封闭的均衡系统。这种做法奠定了今天我们所看到的传统经济学的框架。这种错误归类就像一个紧箍咒，迫使经济学家不得不作出一些极度不现实的假设，并且限制了经济学在实证上的成功。

瓦尔拉斯的教堂之外

圣塔菲会议持续了10天，会议结束后人们感到既兴奋又疲惫。[82]尽管经历了大量激烈的辩论和少数几次冲突，但总体而言会议是在一种强烈的彼此尊重的氛围中结束的。经济学家本身对于传统经济学理论的怀疑更深了，并且对长期存在的经济学问题产生了新的思考。物理学家得以大致了解了一种在本质上令人着迷而又复杂、具有挑战性的现象。会议促成的联系和合作在会后依旧在继续，并且圣塔菲研究所还发布了一个关于经济学的多学科交叉的研究项目，最初由布莱恩·阿瑟和约翰·霍兰德担任主席，由花旗银行提供资金支持。就像从山上滚落的雪球一样，

将经济看作复杂适应系统的理念愈演愈烈。圣塔菲研究所依然是其背后主要的驱动力，多年来举办了各式各样的工作坊和会议，涵盖了从金融到经济不平等的诸多主题，10年后还举行了新的会议。[83]也许更重要的是，如今研究这些问题的研究者分布极其广泛，实际上美国每一个重大的经济部门都至少有一人或两人在研究复杂经济学的某个方面。并且，跨学科合作也获得了良好的发展，如今在一些非经济学专业期刊（如《物理评论快报》《自然》以及《理论生物学杂志》）上看到经济学文章已不足为奇。

本章所呈现的批评并不是为了表达对经济学或经济学家的言语攻击，而是对构成传统经济学的某些具体理论进行批评。经济学就像一个广派教会[(7)]，在经济学领域之内，传统经济学“正教”之外有许多工作要做，并且许多近期工作正在推动扩展传统经济学的边界。我们将在第二部分看到，许多经济学家本身已经离开传统经济学，带头发展复杂经济学的新理论，他们还常常会与来自其他领域的同事携手合作。我既不是在批评经济学想成为科学的愿望，也不是在批评数学在该领域的应用。许多批评家渴望回到瓦尔拉斯理论提出之前的简单时日——当时的经济学是政治和道德哲学的一个分支，人们无须具备高深的数学技能就能成为经济学家。尽管经济学拥有作为一门科学的许多独特特征，但如果缺乏数学的精准和严谨，它就无法创造出在科学上站得住脚的解释。[84]我所提出的问题是：经济学是否使用了正确的数学工具？

这种批评绝不会抹杀经济学在过去一个世纪取得的成就。如前所述，科学理论永远只是近似于它们试图描述的潜在事实。在第二部分，我将论证复杂经济学比传统经济学更接近经济事实，这就好比爱因斯坦的相对论比牛顿三大定律更接近物理事实。尽管如此，我仍然相信传统经济学在许多重要的问题上接近正确或在方向上是正确的，而且它的理念改善了数十亿人的生活。此外，传统经济学的理论对于人们来说具有一种真正的智力美，并且被最聪明的人发展出了科学的优雅。一代人在前人的基础上提出理念，偶尔进行替换，科学就是这样进步的。经济学获得进步，超越瓦尔拉斯教堂建设者所创建的范式——没有什么比这更能为他们的遗产增添荣耀。

第二部分

洞悉财富起源真相的工具：复杂时代需要复杂经济学

THE ORIGIN OF WEALTH

THE ORIGIN OF WEALTH

真正的发现之旅，

不在于寻找新的风景，

而在于拥有一双发现的眼睛。

——马塞尔·普鲁斯特

04

定义复杂经济学：审视经济的五种视角

菲律宾首都马尼拉的周边有一个堪称世界之最的垃圾场，这个垃圾场有一个与它本身极不和谐的名字：乐土。[1]玛丽亚·卢兹·奥孔德拉（Maria Luz Ochondra）就生活在这里。没有人知道乐土及周边生活着多少人，但至少有数千位成人和孩子每天都在这个垃圾堆里寻找有用的东西。神父乔尔·伯纳多（Joel Bernardo）是当地的一位天主教牧师，他说这些人似乎是世界的“垃圾”，穷人中的穷人。

另外一位居民帕斯·卡洛佩兹（Paz Calopez）则这样描述乐土：“垃圾场总在冒烟、冒火，下雨的时候也是一样。垃圾燃烧着，在夜里也不停歇，发出‘噼噼啪啪’的声响，我们觉得那是电池爆炸的声音。垃圾场的气味比厕所还要恶心，尤其是推土机作业的时候。”2000年7月，这里发生了一起事故，一连数周的雨水浸湿、泡软了小山似的垃圾堆，垃圾场突发崩塌，致使200余人丧生，其中就包括奥孔德拉的两个儿子。

我们很难相信，在这个地球上最像地狱的地方，在这个最绝望的人群居住的地方，却萌生了一种复杂的、甚至可以说是生机勃勃的经济形态。1994年，菲律宾政府关停了另外一处大型垃圾场，导致送往乐土的垃圾数量大大增加。乐土不断被填满，此地的经济也几乎在同一时间焕发了生机。处于乐土经济底部的是拾荒者，他们以垃圾为生，每天的工作就是在垃圾堆里寻找废金属、塑料瓶、橡胶轮胎和其他有用的东西。有时他们还能找到废弃的家电、家具、衣物和儿童玩具，甚至还有食物。对于捡来的东西，拾荒者会将其中一部分留作自用，但大部分都会卖给垃圾贩子。

垃圾贩子则是经济链条的下一环节。垃圾贩子会对从拾荒者那儿收来的东西进行分类，然后将金属、塑料和橡胶转卖给回收商和制造商。有些垃圾贩子甚至会同大型公司或连锁酒店签订回收合约。在卡车将垃圾倒到普通垃圾堆之前，合约允许垃圾贩子雇用拾荒者对这些客户的垃圾先扒拉一遍，挑选出有价值的部分。拾荒者是按件收费的。由此，垃圾贩子能更有效地筛选出更有价值的东西，而他们的客户则可以少付一

部分垃圾处理费。

经济链条的再下一环节则是各种针对拾荒者和垃圾贩子及其家人的服务。各色棚屋小店向他们出售各种各样的物品，例如肥皂、鞋、自行车零件、冰激凌和学习用品等。尽管拾荒者的生活非常艰难，既不健康也不安全，但垃圾堆还是为他们提供了相对稳定的收入，使得他们中的大部分人不至于饿死——在一个有7400万人口每天人均收入不足1美元的国家，被饿死是一种真真切切存在的风险。正如神父伯纳多描述的那样，“原始资本主义在这里发生着作用，它确实创造了财富，这里每天可以产生数百万比索”。

经济源自何处？像乐土这样的复杂经济生态是如何从一堆垃圾中产生的？这些行为、关系、组织和理念是如何支撑起一种经济形式的？随着时间的推移，它们又会如何变化？起源问题对于大多数学科而言都是至关重要的。我们很难想象，如果现代宇宙学中没有大爆炸，或者生物学上没有进化论会是什么情形。同样，如果无法回答“经济从何而来”这个问题，我们很难相信其会成为一门真正的学科。然而，经济的起源问题并未在传统经济学中占据重要的地位。[2]如前所述，传统经济学倾向于关注经济蛋糕如何分配的问题，但不关注经济蛋糕最初是如何出现的。传统的增长理论甚至以“假设存在这样一种经济形式”作为开头。经济的形成过程是一个一流的学科问题，并且是传统经济学与我所谓的复杂经济学的一个显著区别。

欢迎来到糖域的世界

乔舒亚·爱泼斯坦（Joshua Epstein）与罗伯特·阿克斯特尔（Robert Axtell）是来自布鲁金斯学会的两位研究者，该学会是华盛顿特区领先的公共政策智库之一。1995年，他们决定开展一个实验，以验证是否可以从零开始构建一种经济形式。就像生物学家试图在培养皿上创造出试管生命一样，爱泼斯坦和阿克斯特尔希望验证能否在电脑模拟的世界里，在硅片上创造出一种经济生命体。[3]

传统的微观经济模型通常以这样一种假设作为开始：消费者、生产者、技术和市场都是存在的；宏观经济学模型则以假设货币、劳动力市

场、资本市场、政府和中央银行的存在作为开始。爱泼斯坦和阿克斯特尔没有假设这些元素，他们希望回到最初的原点，回到一种自然的状态，他们的模型里只存在拥有几项基本技能的人物，以及具备一些自然资源的环境。他们希望找到能够激发经济产生连锁反应的最小条件。那么，我们需要什么条件才能创建一个系统，并使它不断走向更高的经济秩序呢？[4]

为了描述爱泼斯坦和阿克斯特尔的模型，我们可以想象有一群人因为海难来到了一个荒岛上，但荒岛和这些遭遇海难的乘客都是电脑虚拟出来的。电脑里的荒岛是一块方方正正的场地，上面横竖各有50个格子，就像一个巨大的棋盘。这个虚拟的荒岛上只有一种资源——糖，每个格子里糖的数量均有差别。糖堆的高度范围在0（没有糖）到4个糖单元（最高值）之间。糖被堆在两座糖山上，一座在最北端，另一座在最南端，糖山的高度分别为3个单元和4个单元（见图4-1）。两座糖山之间有一块“贫瘠地”，那儿几乎没有糖。爱泼斯坦和阿克斯特尔将他们想象出来的糖岛称为糖域。

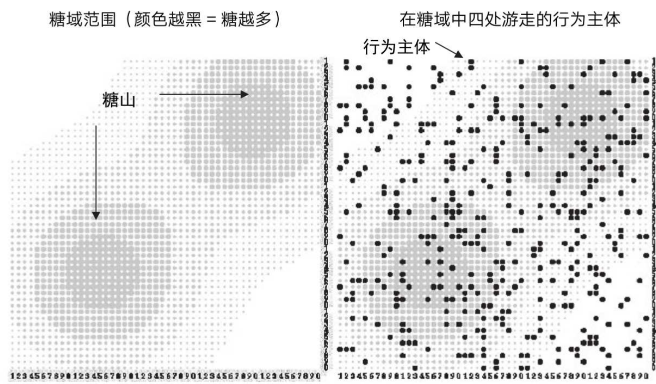


图4-1　糖域

资料来源：爱泼斯坦与阿克斯特尔（1996）。

当然，糖域的地形是真实岛屿的极简版，但它强调了真实岛屿的三个主要特征：（1）存在物理空间的概念。也就是说，在糖域上，你可以向东、南、西、北四个方向移动。（2）存在能源资源：糖。（3）地形地势变化不一，既有高山又有峡谷，既有沃土又有荒地。

同样，虚拟人物是真实人物的极简版，但他们之间有一些关键的共同之处。每一位虚拟人物，或称“行为主体”，都是一个独立的、可以从糖域环境中获取信息的电脑程序，他们可以通过自身代码对信息进行计

算，作出决策、采取行动。在最基础的模拟版本里，糖域上的每一位行为主体只能做三件事情：找糖、移动、吃糖，仅此而已。为了寻找食物，每位行为主体都拥有四处察看的视觉能力，并且拥有朝着糖堆移动的能力。每一位行为主体都可以通过新陈代谢来消化糖。

爱泼斯坦与阿克斯特尔希望验证简单行为主体在简单的环境中能否创造出类似于经济的东西。因此，在游戏的每一个环节，每一位行为主体都必须遵守一套基本的规则：[5]

- 在视野范围内，行为主体可以向坐标的四个方向张望：东、南、西、北（行为主体无法斜视）。
- 行为主体可以判断出在视野范围内，哪一个尚未被占领的方格拥有最多的糖。
- 行为主体可以移动到方格内并吃糖。
- 行为主体吃糖的数量计入贷方，新陈代谢所消耗的糖的数量计入借方。如果行为主体吃的糖的数量比消耗的数量多，那么它存糖账户上的数值就会增加（你可以想象一下身体脂肪的增加），并且这些储蓄会进入下一环节。如果它吃的不如消耗的多，就会花光储蓄（消耗脂肪）。
- 如果行为主体存糖账户上的数值小于零，那么行为主体就会被判定为饿死，被移除出游戏。如果情况相反，行为主体就会一直活到系统设定的最大年纪。

为了开展这些任务，每一位行为主体都有通过遗传获得的视力和新陈代谢能力。也就是说，每一位行为主体都关联着相应的电脑代码，即电脑DNA，这决定着行为主体能看见前面的多少个方格以及每个来回需要消耗多少糖。一位视力良好的行为主体能看见前方的六个糖方格，而视力不佳的行为主体只能看见一个糖方格。同样，新陈代谢慢（好）的行为主体在游戏的每个环节只需要1个糖单元，而新陈代谢快（差）的行为主体则需要4个糖单元。视力和新陈代谢遗传在群体中是随机分布的，因此，行为主体的群体是多种多样的。一些行为主体的视力很差、新陈代谢很好；另一些则视力好、新陈代谢差；还有的视力和新陈代谢都很好，也有倒霉透顶的——两者都很差。行为主体的最大寿命也是随机设定的，到了既定寿命后，电脑死神会将它们从游戏中移除。最终，当糖被吃掉，它会像庄稼一样又在原地长回来，生长速度是每个时间段

生长一个单位。因此，如果一个高度为4的糖堆被吃掉了，它需要4个时间段来恢复到原来的高度。

在游戏一开始，250位行为主体被随机安置在糖域里。某些行为主体恰巧被安置在多糖的山上，生来就有很多糖，而另外一些行为主体运气不佳，出生在少糖的贫穷地区。游戏开始之时，场面显得有点混乱，行为主体们四处奔波寻找糖，许多降生在少糖地区的行为主体被活活饿死。很快糖域中就出现了秩序。就像我们能够预料的那样，有行为主体发现了两座多糖的山，然后开始在山的周围集聚（见图4-2，其中1、2、3、4是连续的时间点）。这样一来，出生地所决定的命运马上就开始起作用。由两座糖山所在的多糖区域所组成的岛结构使得行为主体分成了两组。[6]如果将注意力放在糖域里某个具体的区域，我们就会发现一些其他的东西。行为主体的吞食效率十分惊人，在很长一段时间内，附近的糖几乎所剩无几，即便我们只为行为主体设定了极其简单的规则，他们也会高效地按最大量摄入。当糖堆又长到可以被收取的时候，它几乎立马就会被虎视眈眈的行为主体吞食。因此，他们几乎马上就能看到一个自我管理的基本模式。行为主体快速地组成两个集中的“部落”，分别分布在两座糖山上，然后高效地“收割庄稼”。

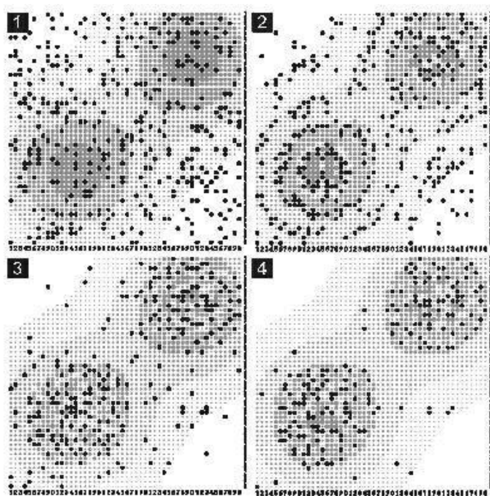


图4-2　糖域上所呈现出来的秩序

资料来源：爱泼斯坦和阿克斯特尔（1996年）。

富者愈富

爱泼斯坦和阿克斯特尔对模拟过程中糖域群体的各种数据进行了收

集和追踪。

他们追踪的其中一个变量是行为主体的财富，即某个时间点上行为主体储蓄账户里糖的数值。在每一个阶段，研究者都会找出最富有的行为主体（比如其财富值为100），然后将最富有与最贫困的行为主体之间的差距分为10个数值段，比如，0～10、11～20、21～30，等等。接下来，他们会计算在每一个数值段里有多少位行为主体（有两位行为主体的财富值为0～10，六位为11～20等）。从这个直方图中，爱泼斯坦和阿克斯特尔得到了一些关于财富随着时间变化的有趣发现（见图4-3）。在模拟实验的开始，糖域是一个非常平等的社会，财富分布十分均匀，钟形曲线上只有少数几位行为主体极其富有，少数几位极其贫困，中产阶级占了大多数。此外，最富有者与最贫困者之间的差距相对较小。

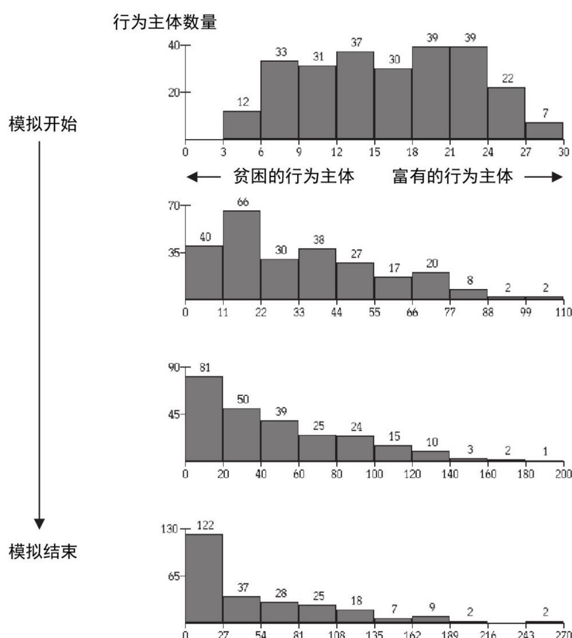


图4-3　糖域中财富分布的进化

资料来源：爱泼斯坦和阿克斯特尔（1996年）。

随着时间的推移，分布情况出现了极大的变化。当行为主体在两个糖山集合时，整个糖域的平均财富增加了，但财富的分配却出现了偏态分布，即出现了一些超级富裕的行为主体，一大群雅皮士行为主体，中产阶级的规模不断缩小，而贫困行为主体的规模不断扩大。如图所示，

坐标轴的规模在随着时间不断变化，最初最富有的行为主体只有30个糖单元，但在最后，最富有的行为主体有270个糖单元。

在第2章中，我们简要讨论了19世纪的经济学家维尔弗雷多·帕累托的工作。除了发展“帕累托最优”的概念，这位意大利经济学家还对社会财富的分配有着极大的兴趣。1895年，他收集了多个国家国民的收入数据，并将这些数据制作成了一个分布曲线图，即“帕累托分布”。帕累托分布并不是一个钟形的正常分布，其主要是指许多人位于财富分布的最底端，中产阶级十分广泛，少数人超级富有。帕累托分布也是80/20原则的出处，意指80%的财富掌握在20%的人手中。在过去一个世纪里，经济学家已经证实，人们的收入和财富大致符合帕累托分布，并且这个结果与不同国家不同时期情况的相符程度非常高。可以说，简单糖域模型中的财富分布如实反映了现实世界的帕累托分布。[7]

有人会问，为什么会发生这种情况呢？为何在糖域中都会富者愈富、贫者愈贫？在糖域这个可控的世界里，验证各种假设是一件非常简单的事情。首先我们可能会问，这是自然现象吗？它是否与参与者的基因遗传有关？是不是所有具有良好视力和缓慢新陈代谢的行为主体都获得了较多的财富？答案是否定的。基因遗传在设定时是均匀随机分布的。如果财富与糖域里的基因遗传有关，那么财富的分布应该十分平均才是，即富有者、中产者和贫困者的数量应该大致相同。其次，有人可能会问，这是由于行为主体的出生环境所导致的吗？换言之，是否所有出生在糖山之巅的行为主体都能够获得大部分财富，而出生在贫瘠之地的行为主体会一直贫困？这个问题的答案也是否定的。同基因遗传一样，行为主体的出生地也是均匀随机的，因此，如果这是决定行为主体最终经济阶层的原因，那么他们在各阶层的分布也应该是均匀的。但为何从最初的随机条件开始，最后得到的却是财富的偏态分布呢？

答案是，本质上说是“因为所有因素”。偏态分布体现了系统的涌现特性[8]，这是行为主体群体的集体微观行为所涌现出来的宏观行为。物理景观、行为主体的基因遗传、他们“出生”的地点、遵守的规则、彼此之间互动的动力以及最重要的运气，所有这一切联合起来导致了财富偏态分布的涌现结果。为了弄清其中的原理，我们可以假设行为主体1和行为主体2都具有中等的基因遗传，并且都出生在中等糖量的地区。在出生之时，他们成功或失败的机会是相等的。但行为主体1在进行生命中的第一次移动时，他向各个方向都看了看，发现其他行为主体已经占领了东、西、南三个方向的方格。于是他向着北方出发，一路吃糖。很

偶然的情况下，北方指向的是一座糖山的中心地带，并且同样很偶然的是，那里的行为主体很少。于是，在接下来的几个回合中，行为主体1尽最大可能大饱口福，累积了大量储蓄，直到其他行为主体发现这个区域并进入。然而，正是这个最初的黄金时间段使得行为主体1的财富远远高于平均值，并且让他作为最先进入糖山的行为主体之一留在那里，舒舒服服地度过余生。行为主体2就没有那么幸运了。第一次移动的时候，他选择了向南，那是一片贫瘠之地。当他意识到自己的错误时，其他行为主体已经占领了富饶的北方地区，这让他无法在几个回合内就回到富庶之地。在做这些动作时，他的储蓄不断减少，离平均值越来越远。当他找到去往更加富庶区域的道路时，那个区域已经人满为患，因此他很少有机会发现没被占领的牧场。

行为主体2存活了下来，但是随着游戏的继续，他与具有同样禀赋的行为主体1的差距越来越大。因此，即便行为主体具有同样的初始禀赋，早期微小的偶发事件也会随着游戏的推进被放大，最终造成两位行为主体截然不同的结局。这种结局就是经济学家所谓的横向不平等，这在传统经济学理论中是被严格禁止的。在传统经济学的均衡世界里，具有同样能力、偏好和禀赋的人们最终应该获得同等水平的财富值，如果出现了差距，那就是随机分布的噪声。但在糖域这个非均衡世界里，横向不平等是一个无法改变的事实。值得一提的，行为主体1和行为主体2之间的结果差异不仅仅是由于系统的随机性（因此不是传统经济学所谓的噪声）。随机性起到了让行为主体选择不同的道路的作用，但游戏的动态使得道路与道路之间的分歧越来越大，最终结果的偏态程度也就比随机因素单独导致的更加严重。

为了弄清动态是如何发挥作用的，我们可以假设行为主体1的基因相对更优，或者出生地点的情况相对更好，那么他就能走得更快、更远。同样，假设行为主体2的初始禀赋稍差，他每况愈下的程度更加严重，这就很有可能导致他食不果腹。这样一来，个体能力、出生环境以及个体命运的迂回曲折，所有这些加在一起，就为每一位行为主体的模拟人生创造出了独特的道路。关键之处在于，某一个时间点上的小小区别（比如在某处的幸运或倒霉际遇）就会导致道路尽头结局的重大差别。这些小小区别的加速会让一些行为主体变得富有，让另一些行为主体变得潦倒。此处可以想象一下模拟开始时我们提到的钟形曲线，现在可以将曲线两端的尾部向外拉伸。右边的尾部延伸至将一些超级富豪包含在内，而中产部分凸起的形状会收缩。当我们伸展左边（贫困）的尾部，又不能让他们的储蓄比糖值为零更低（在零的时候行为主体就会死

掉），这样一来这部分的尾部就会被砍掉。延伸的、被砍掉的分布会让我们得到图4-3底部的结果。

所有这些意味着哪怕是在糖域中，也不存在导致贫困和不平等的简单因果关系。相反，它是一个复杂的、各种因素的集结。哪怕是在糖域这样高度简化的世界里，找到贫困问题的解决办法也非易事。爱泼斯坦和阿克斯特尔指出，糖域太过简单，以至于无法得出关于现实世界贫穷和不平等的具体结论。但模型显示，单一维度的观点——无论是左派观点（例如，贫困是由于富人剥削穷人）还是右派观点（例如，如果你很穷，那肯定是因为你愚蠢或者懒惰，或者又蠢又懒）都有可能是错误的，并且任何有效的解决方案都必须要系统地解决问题。复杂经济学家超越了糖域的简单世界，可以让我们更好地理解现实世界中的不平等，我们将在第18章看到他们的工作成果。

小鸟这样做，蜜蜂这样做……行为主体也这样做

爱泼斯坦和阿克斯特尔在糖域中加入了另外一个因素，他们知道，在现实世界里，这个因素能把世界搅得天翻地覆，那就是——性。经济学家已经证明，性对于经济学的影响至少是从19世纪初期托马斯·罗伯特·马尔萨斯写出著名的《人口论》时就已经被发现的。幸运的是，像糖域这样的电脑模型能让我们实时地观察人口增长和进化的作用，并通过各种条件和假设进行实验。

爱泼斯坦和阿克斯特尔决定为每一位行为主体贴上一个注明年龄和性别的标签。[8]当行为主体达到“生育年龄”，并且拥有最低限额的糖储蓄时，他或她就被认为是有生育能力的。在每个阶段，能够生育的行为主体会全面扫描所在方格的左邻右舍。如果发现异性的适育行为主体，他们就会配对并进行繁殖。新生儿的DNA一半来自母亲，一半来自父亲。因此，孩子的视力和新陈代谢能力是父母亲能力的混合体。此外，孩子还能够继承父母亲的财富，他们的财富值等于一半父亲的财富加上一半母亲的财富。婴儿行为主体的出生地是邻近父母亲所在方格的一个空方格，因此孩子所在区域的糖富有量与父母所在区域的类似。这对于高度简化的设置有两个重要作用：一是群体中健康状况最差的成员拥有后代的可能性最小；二是健康状况好的成员的后代在起跑线上拥有明显的生理和环境优势。

当爱泼斯坦和阿克斯特尔在新增了性因素的糖域按下“运行”按钮

时，行为主体开始像往常一样匆忙奔走，收糖。适育的行为主体很快就能配对、生子，春去夏来，糖域上随之出现小小行为主体“啪嗒啪嗒”的脚步声。一段时间后，爱泼斯坦和阿克斯特尔发现了三件事。

第一，群体中最不健康的成员会消失，而最健康成员的后代越来越多。随着时间的推移，行为主体的视力和新陈代谢效率的平均值开始提升，这些因素的提升又会导致平均财富值的增加。

第二，新的出生-死亡动力将会引发人口波动。在性因素出现之前，行为主体数量与土地的承载能力大致相当。然而，丰收和饥馑的循环伴随着性出现了。健康的行为主体能够进行储蓄，拥有很多子孙，但最终，人口的增长会超过土地的承载能力，于是行为主体开始对糖储备进行过度放牧，引发饥荒，导致人口数量急剧下降。最后，土地再生，循环又从头开始往复。

第三，贫富差距越来越大。在第一部分，我们已经看到，没有性因素的简单模型里也能产生偏态财富分布。代际基因遗传和财富继承的引入进一步加速了富者愈富、贫者愈贫的趋势。

伸进糖域的“看不见的手”

至此，糖域里的行为主体已经成为纯粹的狩猎采集者，他们收集和消费从糖域里发现的糖。但爱泼斯坦和阿克斯特尔又在这个人为创造的世界里加入了更多现实主义的因素。[9]他们引入了第二件商品——香料，每个方格的价值由糖和香料的储存量共同决定。跟糖一样，香料集中在两座山上。糖山分布在东北和西南，而香料山分布在东南和西北。爱泼斯坦和阿克斯特尔还稍微调整了行为主体的新陈代谢，让所有的行为主体都需要两种商品才能存活。其中有所区别的是，部分行为主体需要大量糖和少量香料，而另一部分行为主体需要大量香料和少量糖，这些是由他们的DNA决定的。在传统经济学的术语中，这种需求的差异可以被认为是行为主体的偏好。[10]跟之前一样，行为主体可以储蓄他们消耗不完的糖和香料。

在最后一个步骤中，爱泼斯坦和阿克斯特尔对行为主体开放了交易权限。这里没有传统经济学模型中关于市场或拍卖师的假设，只有个体之间直截了当的以货易货。行为主体在糖域内活动时会遇到其他行为主体，在游戏的每个回合，每位行为主体都会朝东、南、西、北四个方向的方格张望，并询问邻近区域的其他行为主体是否愿意交易。如果某位

行为主体拥有许多香料同时需要获得糖，而另外一位行为主体的情况正好相反，那么两位行为主体就可以通过交易来改善彼此的状况。当双方同意进行交易，他们就会进入一个议价环节，向对方出价，直到在价格上达成一致。糖域里没有钱的概念，因而价格就代表着糖相对香料的价值，或者香料相对糖的价值。

爱泼斯坦和阿克斯特尔在糖域里安排了拥有随机天赋、当下可以进行交易的行为主体，然后按下了启动按钮。行为主体开始四处活动，很快一场关于糖和香料的交易就达成了。就很多方面而言，糖域中的交易模式与传统经济学预测的非常相似。爱泼斯坦和阿克斯特尔将允许行为主体进行交易的模式与行为主体自主寻找糖和香料的模式进行了对比，结果发现，交易让糖域社会变得更加富有了。通过交易实现财富增长是经济学的基本预测之一，这可以追溯到亚当·斯密时期。爱泼斯坦和阿克斯特尔用一个极端的案例来解释这种作用：假设有一对邻近的行为主体，其中一位拥有糖，却因为没有香料而濒临死亡；另外一位拥有香料，却因为没有糖而气息奄奄。如果我们禁止交易，两位行为主体都会死；如果允许交易，两位都能活下来。交易的作用相当于提高了土地的承载能力，从而改善了所有行为主体的处境。

爱泼斯坦和阿克斯特尔还跟踪了交易网络的发展，主要是谁跟谁交易，结果发现，不出所料，交易网络中出现了一些地理区域的集群。行为主体们往往与本区域的行为主体交易最频繁，并且由于交易能够带来自我强化，所以一场交易又会引发下一场交易。集中交易网络集群的建立，几乎就像本地市场的城镇。此外，第二种商品的加入让行为主体的移动变得更加复杂。他们不能再简单地与族群一起集聚在两座糖山的顶峰。行为主体们还需要出去寻找香料和交易伙伴。行为主体们在糖山和香料山之间来回穿梭，地理和人口动态叠加在一起，创造出了交通繁忙的贸易路线，这就好比电脑版的古代丝绸之路。

爱泼斯坦和阿克斯特尔在游戏的每个阶段都能看到每位行为主体的内心活动，判断出他们在面对不同的价格时分别愿意买卖多少糖和香料。所有行为主体的这些数据加起来就形成了糖域里的糖和香料的两条供需曲线图。其中一条是像教科书般由左至右下斜的需求曲线，另一条是从左至右向上扬起的供应曲线，爱泼斯坦和阿克斯特尔在他们的模型中并没有加入任何关于供需的具体设置（见图4-4）。这些曲线纯粹是行为主体群体在简单的互动中呈现出来的自下而上的现象。

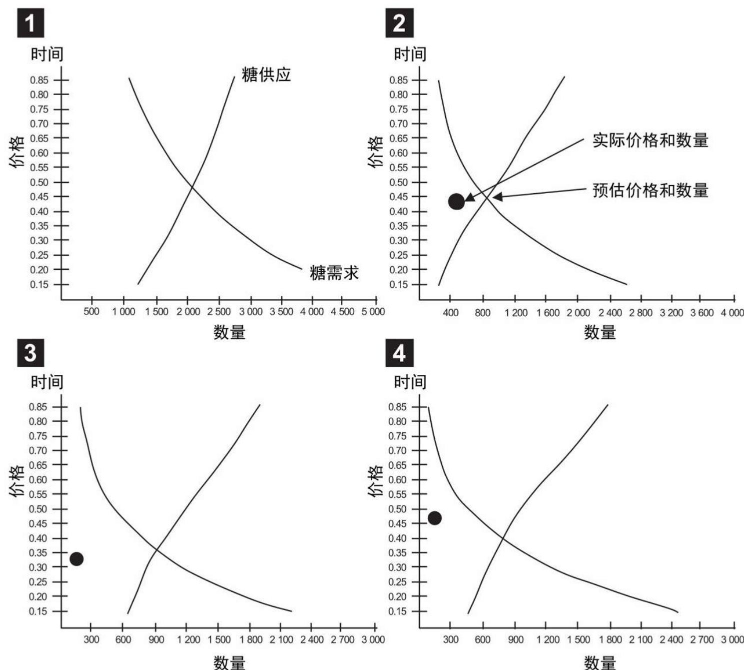


图4-4　糖域中并无均衡

资料来源：爱泼斯坦和阿克斯特尔（1996年）。

遗失的均衡

在第3章中，我们曾讨论过有许多传统经济学的核心预测乍一看来很有道理，但经不住仔细推敲。糖域也存在类似的情况。糖域经济自然而然地产生了X形的供应曲线，实际价格和交易数量（由图4-4中的点表示）从未达到理论预测的均衡点（即图中供应和需求X形曲线中相交的点）。价格会在均衡点附近波动，在某些情况下，甚至会朝着它不断地收敛。然而，哪怕在长期的模型中，价格与均衡之间的偏差依然十分明显。传统经济学或许会回应说，这只不过是“均衡周围的噪音”，但我们应该注意到，本模型中没有任何噪音。初始状况是随机设置的，一旦模型开始运行，正如计算机所设定的那样，所有行为都是完全确定的。正确的解读应该是，糖域中的价格围绕着“吸引子”波动，但永远不会真正地停在均衡点上。因此，在糖域中，正如在现实世界中一样，供求定律仅仅是一个大致的近似状况。

爱泼斯坦和阿克斯特尔还发现，如果系统要想达到均衡，需要比预测的多得多的交易量。传统的关于均衡价格的推论认为，只需要最少数

量的交易就能使供需达到平衡。传统经济学中一个重要的未解之谜在于，为何产品市场和金融市场中的交易量要远远高于理论的预测。[11]同样的谜在糖域中也出现了——价格波动和交易量比传统理论预测的要高。传统经济学模型简单地假设，所有人口只会进行达到均衡所需的最低程度的交易。糖域中不存在协调整个群体的机制，并且由于行为主体被物理距离隔开（因而需要时间来移动），行为主体只能与同一区域的其他行为主体进行交易，永远不能达到全域的均衡。因此，随着行为主体的移动和偶然相遇，新的交易机会就会产生，从而成就了大量的交易。

正如之前讨论的那样，传统经济学的另一个主要预测是一价定律：在特定的市场中，某样商品应该只能以均衡价格进行交易。然而，在糖域中，任何一个时间点上的价格都有很大的差异——在糖域的某些区域，你很有可能被宰，而在其他区域则可以议价，这就像伦敦的番茄酱市场。这是因为，均衡模型假设所有事情是同时发生的，与此不同的是，糖域的情况在随着时间的推移而逐渐明朗，并且行为主体需要时间来寻找交易伙伴。如果在世界的某个区域存在便宜的商品，行为主体需要时间来找到这些区域并移动到那儿，这就好比经济学家在街上发现20美元的钞票一样。当套利的行为主体完成一笔交易，另外一笔交易就会在别处突然发生，变化太多，以至于行为主体永远不可能跟得上经济的不断变化，因此永远不能驱使价格完全达到均衡。

传统经济学的另外一个基本原则是帕累托最优——市场永远会获得资源完美分配的结果；如果重新进行资源分配，要想使部分人的利益有所提升，势必要损害另一部分人的利益。然而，糖域市场的运行并未达到帕累托最优。有些交易本可以达成，本可以让人们的利益得到提升，事实上却没有。跟前面一样，这是因为行为主体处于不同的时间和空间，因此位于西南角的行为主体A不知道在东北角的行为主体B是一个很好的交易伙伴。此外，即便他们知道了，也需要时间才能会合（到那时，价格可能已经发生了变化）。传统经济学家会立即指出，这正是市场解决的问题——市场可以让不同时间和空间的买方和卖方相遇，从而让各方的交易更加高效。我们可以轻易地提升糖域的经济效率——让所有行为主体在下午几时到某个地点集合，然后按照瓦尔拉斯模型那样进行糖和香料的拍卖。

传统经济学假设现实世界是由随时随地发生的“瓦尔拉斯拍卖”构成的，但其实在现实世界中并非如此。在金融市场和eBay网站之外，现实

世界中的拍卖其实比较罕见。为什么呢？尽管拍卖十分有利于交易双方发现“对的价格”，但它处理交易的过程非常缓慢（在eBay上购买某样东西通常需要好几个小时，甚至好几天）。现实世界中的大部分交易与糖域中的情况更为接近——交易是双边的，也就是说在两方之间进行，例如你和街角的杂货店老板，或者IBM公司与福特公司。虽然双边交易产生了价格差异，但是时间效率更高（更确切地说是从计算的角度来说效率更高）。试想一下，如果你每次去本地的超市都要坐在那儿经历一场拍卖，哪怕你只是想买一盒牛奶，这该多么烦人。[12]就像糖域一样，现实世界中的双边交易是在不同的空间和时间进行的，因此任何形式的全球均衡都很难实现。传统经济学假设交易以拍卖方式进行的唯一原因在于，这样就能在数学上找到一个均衡的结果。

尽管糖域里的交易能够“普渡众生”，使得整个社会更加富裕，但它也进一步加剧了贫富差距。在模型的基础版本中，我们从游戏的动态里看到了很小但很重要的贫富差距。在引入性因素和家庭因素后，行为主体的能力和财富可以遗传给下一代，从而导致了贫富差距被进一步扩大。最终，交易的引入再次扩大了差距（虽然每个人都明显变得更加富有了）。爱泼斯坦和阿克斯特尔指出，差距的扩大甚至达到了这样的程度：财富的偏态分布开始接近现代现实世界的经济状况。[13]

等级的进化

迄今为止，糖域的经济只是个体的集合——尚未出现老板和工人的等级，或是供应商、中间商和零售商。这是因为行为主体之间的互动是短暂的，并且是对称的。然而，等级是现实世界经济的标志之一。爱泼斯坦和阿克斯特尔对行为主体的行为做了一个简单的改动：允许进行借贷，这个改动导致交易关系获得了稳定的结构，导致了等级的出现。

在糖域中，成为借款者的原因只有一个——生孩子。糖域里所有行为主体都拥有生育的基本欲望，但有些行为主体太穷了，养不起后代，而另外一些行为主体的糖和香料有富余，无处消耗。因此，爱泼斯坦和阿克斯特尔引入了一条规则，即如果某位行为主体的年纪太大而不能再生育，或者除了自身的生育需求外还有富余的储蓄，那么他就能成为贷方。反过来，如果某位行为主体的储蓄不足以生育后代，但拥有持续的糖和香料收入，那么他就能成为借方。为了简化起见，爱泼斯坦和阿克斯特尔将利率和借款周期设定为固定值，并且针对借方违约设定了基本

的条款。因此，任何有信誉的、需要资源来生育的行为主体都能自由地从拥有富余资源的邻居那儿借贷。

让爱泼斯坦和阿克斯特尔感到惊讶的不是借贷这件重要的事情发生了，而是一种复杂的、等级制的资本市场形成了。经过追踪借方和贷方的关系，他们发现部分行为主体同时变成了借方和贷方，实际上就是中间商。糖域里居然进化出了银行！随着模拟的进一步发展，事情变得更加复杂了。部分十分富有的行为主体成了批发商，他们开始向中间商放贷，中间商又向最终的贷方放贷。在某些模拟状况中，等级链条有五层之多。模拟不仅进化出了银行，而且还进化出了糖域版的机构投资者、投资银行、商业银行和零售银行。

与糖域中的其他新兴模式一样，这些信用网络的进化完全不是从上到下生成的。相反，这些大规模的宏观模型是自下而上、自局部地区细微假设的动态互动而生成的。

硅片上的经济体

如果科学家要对两个理论进行比较，他们会依照对应原理。根据对应原理，新理论应该复制旧理论的成功，解释旧理论的失败，并补充旧理论未能提出的洞见。无论是爱泼斯坦还是阿克斯特尔都没有声称糖域本身代表了一种成熟的经济学理论，他们也不会声称这个模型完全回答了“经济的起源”问题。然而，糖域确实遵循了对应原理，并且为我们指出了新的有趣的方向。糖域模型复制了许多传统经济学中基本的、常识性的元素，其中供求定律达到了一种近似状况（现实世界中也是如此），其中的交易也带来了明显的收益。如果模型没有获得这些成果，我们将很难相信它。爱泼斯坦和阿克斯特尔复制了这些经典的成果，但他们的模型却不包含传统经济学模型中典型的、不现实的假设。他们没有假设模型中的行为主体拥有理性这一非凡的能力，没有假设模型中预先存在社会或经济结构，也没有假设一切事情都会瞬间发生。

但或许更有趣的是，糖域还复制了传统经济学中一些主要的反常现象，包括违背一价定律，横向不平等的存在，以及高于传统经济学预测结果的价格波动和交易量。这是因为爱泼斯坦和阿克斯特尔所采取的方法在根本上不同于传统方法论。他们并未假设糖域是一个均衡系统。相反，他们为糖域设计了行为主体与地形的基本结构，但设计之后，他们让糖域自行进化，看看结果到底如何。糖域有可能会达到均衡，也有可

能不会。最终，糖域自然而然地进化出了复杂的秩序、结构和多样性，包括（加上一点想象可以解读为）部落、贸易路线、资本市场，没有任何一项是事先设计好的。它们都是自下而上发展起来的，是从爱泼斯坦和阿克斯特尔为系统设置的简单初始规则发展起来的。

进一步定义复杂经济学

糖域为经济学真正的新方法提供了例证。尽管像糖域这样的模型是相对较新的发展成果（主要是因为它们依赖于快速计算机），但它们是建立在悠长而丰富的思想史之上的。其中著名的历史人物包括博弈论和元胞自动机的发明者约翰·冯·诺依曼（John von Neumann），“奥地利学派”的成员弗里德里希·哈耶克，行为经济学家赫伯特·西蒙和丹尼尔·卡尼曼，制度经济学家道格拉斯·诺斯（Douglass North），进化经济学家理查德·纳尔逊和悉尼·温特，政治经济学家罗伯特·阿克塞尔罗德（Robert Axelrod）和托马斯·谢林（Thomas Schelling），以及计算机科学家约翰·霍兰德和克里斯多弗·兰顿（Christopher Langton）。[14]

在第二部分的这些章节中，我们将讨论这些研究者和其他许多研究者的工作，看看他们是如何一起为该领域构建了一种新的范式。为了描述这项工作的各个方面，经济学家使用了许多术语，例如计算经济学、基于行为主体的建模、社会动力学、进化经济学、行为博弈论、圣塔菲研究所和交互经济学。[15]所有这些术语所指代的内容都各有特色，如果只选择其中一个，恐怕我对该领域状况的审视将会过于狭窄。因此，我选择了一个涵盖性的术语：复杂经济学。如前所述，在这个阶段，复杂经济更像一个研究项目，而非单一、综合的理论，至于这个涵盖性术语应该包含什么内容，其中尚有许多灰色区域。表4-1中描述了五个“重大理念”，它们能够帮助我们定义复杂经济学，并将它与之前的工作区别开来。

表4-1　区分复杂经济学与传统经济学的五个“重大理念”

	复杂经济学	传统经济学
动态	开放的、动态的非线性系统，远离均衡状态	封闭的、静止的线性系统，处于均衡状态
主体	单独建模，使用归纳性的经验法则来做决策，拥有不完整信息，受限于错误和偏差，随着时间的推移，能够学习和适应	集体建模，使用复杂的、演绎性的计算来做决策，拥有完整信息，不会犯错、没有偏差，没有学习和适应的需求（已然完美）
网络	清晰地模拟个体行为主体之间的互动，关系网络随着时间的变化而变化	假设行为主体只通过市场机制进行直接互动（例如拍卖）
涌现	微观经济与宏观经济没有区别，宏观模型是微观层面行为和互动的结果	微观经济和宏观经济是两个互不相干的学科
进化	变异、选择和放大的进化过程使得系统能够创新。进化过程引发了秩序和复杂程度的变化	没有内在性创新的机制，秩序和复杂程度也不会发生变化

复杂经济学方法的一个重要方面在于，这五个领域不能仅用数学定理进行分析。传统经济学依赖于数学证明，布莱恩·阿瑟认为它受限于“羽毛笔和羊皮纸技术”。传统经济学使用了更加广泛的工具包，定理、均衡分析、博弈论和其他传统方法都是工具包的组成部分，但复杂经济学的研究者还充分利用了摩尔定律，从而广泛地将计算能力应用到了工作当中。此外，研究者们还从物理学、生物学和其他领域引入了新的数学工具和统计工具，以帮助他们更好地理解经济作为一个开放动态系统的属性。最终，复杂经济学家利用实验经济学以及经济学数据分析的进步来为他们的理论构建了实验证据主体。

在接下来的5章中，我们将更深入地了解这些“重大理念”。在第三部分，我们将看到它们是如何共同打造基础，让我们得以用全新的视角看待财富的起源的。

05

视角一：
经济是一个动态系统

20世纪伊始，经济学从与物理学的邂逅中抽身，退回了智识上的古巴，物理科学的发展方向发生了根本性转变。在接下来的100年里，物理学家拆毁了瓦尔拉斯及其同伴所借用的那些理论，并以相对论、量子力学、非均衡系统热力学、混沌理论和复杂理论取而代之。科学家开始承认，宇宙既不像钟表一样确定，也不像赌场一样随机。实际上，完全确定或完全随机的系统非常罕见。宇宙中的大多数现象处于两种状况之间：它们用复杂而不可预测的方式将确定性和随机性混合在一起。20世纪，科学开始接受这种混乱的、中间的状况。

我在第3章指出，经济曾被传统经济学错误地划归为均衡系统，但实际上它是一个复杂适应系统。在第4章，通过糖域的案例，我们开始探索复杂适应系统的本质和行为。在本章中，我们将继续完善对于经济的定义以及明确其新特征所代表的含义。自20世纪80年代以来，随着詹姆斯·格莱克（James Gleick）的作品《混沌》（*Chaos*）乃至科幻小说等畅销书以及好莱坞电影的宣传，诸如混沌、动态和非线性等概念进入了大众的认知视野。我们将了解这些术语在经济学环境中代表什么意思，并讨论从飞机制造业到房地产行业的繁荣与萧条等现象与它们之间存在什么关系。

动力学与反馈

观察经济的一个有用切入点是动态系统。也就是说，经济会随着时间的变化而变化。这是显而易见的：价格会上涨下降，工资变来变去，公司进入市场又退出。这也是传统经济学已经承认的动态性，但如前所述，它们通常被视为是由外源性资源所导致的，例如技术变革、政治事件、消费者偏好的变化。我们感兴趣的问题在于，这样的动态行为如何才能从内部产生，即作为经济结构本身的结果而产生？

如果科学家指出某个系统是动态的，他们的意思就是说系统的当前状态是前一个时刻的状态以及两个时刻之间的变化的函数。[1]动态系统的一个简单案例就是银行账户。[2]账户的状态或余额会随着时间变化而变化。明日的余额取决于今日的余额再加上中间一天的变化，例如存款、取款或利息支付。我们可以用一个简单的公式来描述：$B_{t+1}=Bt+$（存款-取款+利息）。当天的账户余额用B表示，t+1代表明天。因此，储蓄账户会随着时间的变化而不断变化。在迭代循环中，某段时间内方程的输出数据会成为下一个阶段的输入数据。动态过程中的变化既可以是不连续的，就像银行账户在特定时间发生的变化（例如在某一天付息）；又可以是连续而平稳的，好比星球的轨道。

要想便捷地描述动态系统，可以使用存量和流量这两个术语。[3]存量是指事物的累积，例如银行里的存款余额或浴缸里的水。存量随着时间变化而变化的速率就是流量，例如钱流入或流出银行的速率，或者水流入或流出浴缸的速率。经济是由各种各样随时变化的存量组成的，比如货币总供应量或就业人口数。其中的每一种存量都有对应的流量或变化率，例如，中央银行也许会增加或减少货币供应，公司或许会雇用或解雇员工。需要注意的是，流量总是以某个时间单位为单位的。存量和流量不一定是指有形物，例如货币和人，也可以指较难感触到但依然非常重要的因素。比如消费者的信心就可以被视为一种随着时间的推移而不断提升或下降的存量。[4]

如果将经济看作各种存量与相关流量的集合，我们很快就能明显地看到，存量和流量是通过复杂的方式相互联系的。例如，如果就业的存量跌至很低，政策制定者或许会决定降低利率以鼓励贷款，这样一来就会增加可以用来投资的货币存量，它们可以被各种公司用来投资新的产能、创造更多的岗位需求，从而提升就业的存量，这最终又会影响未来的利率政策。在动态系统中，存量与流量之间诸如此类的关系链条被称为反馈回路。

如果某个系统的输出数据是另外一个系统的输入数据，那么这两者之间就会形成反馈回路，比如A影响B，B影响C，C又会影响A。如果联系不断增强，可称其为正反馈——我推A，A用更大的力气推B，B用更大的力气推C……如此循环。我们之前也提到过一个正反馈的经典案例：麦克风距离扩音器太近就会造成刺耳的噪音。除此之外，螺旋式下降也是正反馈的一种。举例来说，消费者信心下降会导致消费缩减，从而导致生产缩减，进而导致就业率降低，最终导致消费者信心进一步下

降，消费进一步缩减，螺旋下降直至经济萧条。1936年，约翰·梅纳德·凯恩斯在其著名的作品《就业、利息和货币通论》中就描述了这种动态循环。[5]即便对于人们来说它并没有积极的意义，但这也是正反馈的一个案例。关键要记住的是，正反馈会强化、加速或放大正在发生的事情，无论是良性循环还是螺旋式下降。拥有正反馈的系统因此能够呈现出指数级增长、指数级崩溃或振幅不断扩大的震动。

与正反馈相反的则是负反馈。负反馈是一种抑制性的循环——相对于强化，它的作用是相反的。正反馈会加速变化，负反馈则会抑制变化，使事物重新回到常态。之前我们也提到过一个关于负反馈的经典案例——恒温器。拥有负反馈的系统倾向于把事物推回到某个设定的点、某种均衡或以越来越小的振幅进行震动并逐渐停息。[6]

动态系统还有一个组成部分——时间延迟。在不熟悉的地方（例如宾馆）冲澡时，你大概有过这样的经历：开了热水，发现不够热，于是多开一点，水就变得极烫，然后你就又调小一点，依然很烫，继续调小，结果又太冷了，如此反复。问题在于你对开关旋钮的调节与淋浴器的温度之间存在一个小小的时间延迟。这个延迟导致你调试的水过热或过冷。最终，当你发现问题后，偏差会变得越来越小，直到调出合适的温度。时间延迟越长，你就越难掌握淋浴器，偏差也就越大。

如果一个动态系统拥有多种存量和流量，它们既通过正反馈也通过负反馈进行互动，那么我们不难看到系统将很快变得复杂。正反馈可以驱动系统，使之加速，与此同时负反馈则会抑制系统，削弱和控制它。当加入时间延迟的因素后，驱动和削弱将会脱离平衡，脱离同步，导致系统以高度复杂的方式产生振荡。

非线性系统

我们对于经济的第二个观察结果是：它是一个非线性系统。这个术语通常会引起混淆（甚至是在经济学家中）。由于“非线性”的字面意思是“不是直线”，人们有时就会认为它意味着任何能够产生曲线的功能都是非线性的。在静态系统中，这种理解是正确的。方程$y=mx+b$是线性的，如果用图形来表示，我们就会得到一条直线。另一方面，方程$y=x^2$是非线性的，如果用图像表示，我们就会得到一条指数级曲线。有一个很好的经验法则是说，如果方程的右边全是加、减、乘、除，它就是线性的；如果方程里有乘方、正弦、余弦或其他奇异的东西，那它就是非

线性的。

然而，在谈论动态系统时，我们需要更加谨慎一些，因为如果按照时间的发展绘制图形，线性动态系统会表现出曲线形的行为。前文中提到过银行账户是动态系统。我们可以创造一个简单的线性方程来计算账户余额是如何随着利息的增加而增加的。假设利率是10%（并且我们不会存款或取款），那么关于余额（B）的方程则是$B_{t+1}=B_t(1+0.10)$。如果当前的余额是100，那么在下一个时间段就会是110。这种关系也是线性的，因为方程的右边只有加法和乘法。但是，如果绘制余额（B）随着时间（t）的变化图，我们就能得到一条指数级曲线（见图5-la）。即便随着时间的延伸会出现一条曲线，方程依然被认为是线性的，因为变化率是线性的——在本案例中，变化率是一个常量10%。如果绘制下一阶段（t + 1）比当前阶段（t）的图像（见图5-1b），就可以看清这一点，这就是所谓的蜘蛛网图（cobweb diagram）。解读蜘蛛网图的方法是：x轴代表的是当前的余额，而y轴代表的是下一阶段的余额。它能帮助我们看到变化率是在加速、减缓，还是保持不变。在图5-1b中，我们看到的是一条笔直的线，这就意味着变化率是常量，并且系统是线性的。线性开放系统会表现出一系列随着时间的变化而产生的行为，包括静止、直线式增长或衰减以及指数级增长或衰减——在两个案例中，变化率都是线性的。

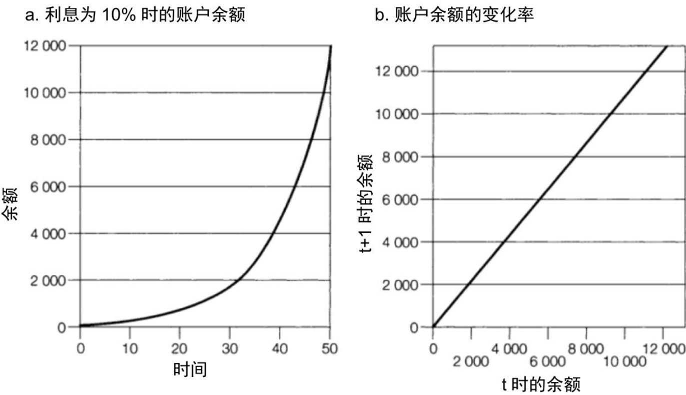


图5-1　指数级增长的账户余额

如果我们观察非线性动态系统，情况会变得更加有趣。还是回到银行账户的案例上来，我们可以想象一个有点奇怪的账户余额是这样的：$B_{t+1} = rB_t-rB_t^2$，其中r是一个常量。等号右边的平方意味着这个方程式是非线性，我们也可以将这个方程改写成$B_{t+1}=rB_t(1-B_t)$，这是大家

熟知的二次映射方程。举个例子来说，如果当前的余额是100，r等于0.10，那么下一阶段的余额将会是-990。虽然这个非线性方程并没有让这个储蓄账户令人满意，但它包含着一些有趣的属性。我们可以将r当作油门，它可以增加或抑制变化的速度，最终，它会让方程呈现出一些与众不同的行为方式。

比如，我们将初始余额设定为0.1，将r设定为1.5，余额将增加到0.3333，然后永远停留在那儿不动（见图5-2a）。如果我们看一看蜘蛛网图（见图5-2b），就能看到变化率随着平缓的曲线不断上升，直到余额达到0.3333，它就不再变化了。这就是我们的老朋友均衡，或者按照动态系统的说法，它是一个定点吸引子——它之所以得名，是因为在蜘蛛网图中，系统会被拉到或者吸引到某个固定的点上（在本例中，这个点是0.3333）。如果我们将r设置为3.3，就会得到一个很不一样的结果。系统会进行规则的振荡，宛如一个来回摆动的钟摆（这被称为周期极限环，见图5-3）。

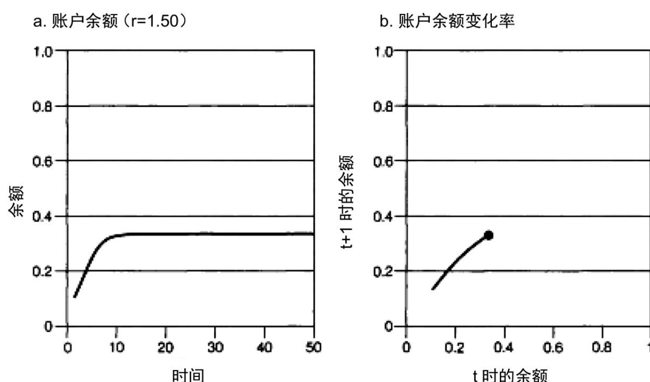


图5-2　定点吸引子账户

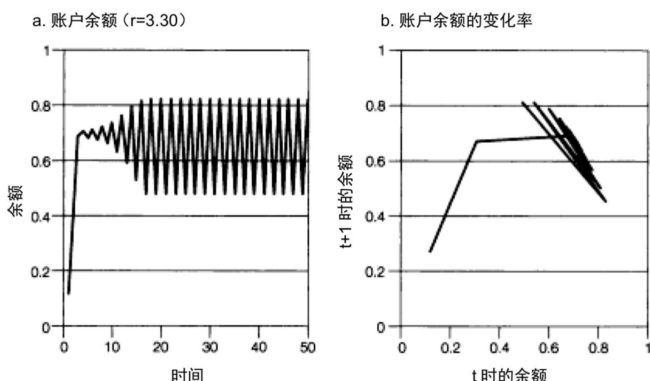


图5-3　周期极限环账户

如果我们将r值稍微调高至3.52，简单的振荡就会变成更加复杂的振荡，就好像心跳模式（准周期极限环，见图5-4）。有一些非线性系统会表现出极其复杂的模式并持续很长一段时间，但最终还是会变成重复。如果我们将r值升为4，就会得到一团混沌（见图5-5a）。混沌系统有三个重要特征。第一，尽管它或许看似随机，但其实它是确定的。举个例子来说，我们的混沌银行账户公式中并没有随机因素。第二，与周期系统不同的是，你可以永远将公式运行下去，并且永远不会重复（尽管有时我们很难将真正的混沌系统与周期很长、很复杂的振荡系统区分开来）。第三，混沌系统是有界限的。即便系统的路线分布广泛，但总有它到不了的地方。在账户余额的案例中，这个值处于0～1之间，而在蜘蛛网图中，我们可以看到，系统的轨迹以一个大致三角形（中间有个洞）的形状向外延伸（见图5-5b）。你会发现要将左边余额随时间变化的图形与右边展示变化率的蜘蛛网图联系起来有点儿棘手，其实并不只有你会这样觉得。非线性动态系统并不都是直观的，我们将会在后面看到，当人们试图在这样的系统中做决策时，这会带来一些麻烦。

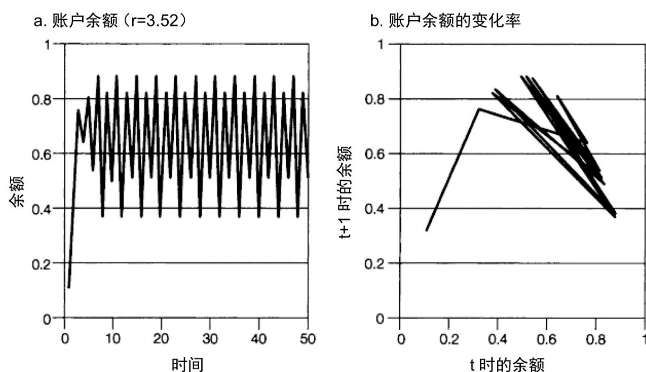


图5-4　准周期极限环账户

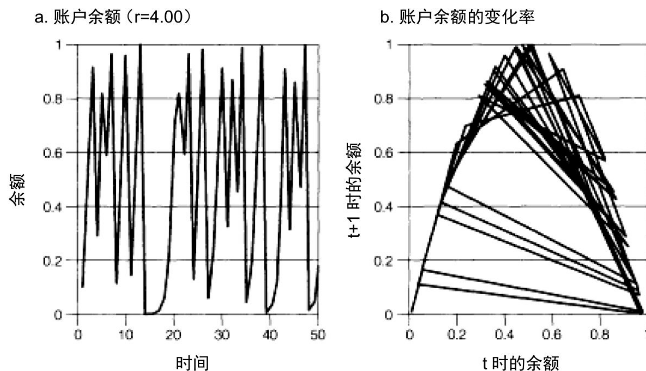


图5-5　混沌账户

仅仅改变一个变量就能使结果大不相同，这表现了非线性动态系统的一个特性：对于初始状况的敏感性。你可以想象一位高尔夫球员在一片高低起伏的草地上击球。如果球员进行了两次轻击，但是球的起点位置或挥杆的角度和力量稍有不同，两个球的走向就会大不一样，最终在草地上的位置也相距甚远。非线性会使得初始时的一点小差别随着时间的推移而被放大，因此，除非你无限精准地了解系统的初始状态，否则就无法知道最终的状态。

相关的特性是，非线性动态系统具有路径依赖性，换言之就是其会受到历史影响。t ＋ 1时的银行余额取决于t时的余额，同样，高尔夫球手的第四杆取决于第三杆，第三杆又取决于第二杆，以此类推。在这一连串的事件中，发生任何一点变化都会使得结果大不相同，比如，第二杆如果打得不一样就会彻底影响到第四杆。

这两个特性，即对于初始状况的敏感性和路径依赖性，使得非线性系统难以掌控，并且在许多情况下都无法预测。事实上，差不多就在瓦尔拉斯思考物理教科书的那段时间，法国数学家亨利·庞加莱（Henri Poincaré）发现了混沌，并且证明许多类型的非线性动态系统可以用当时的数学工具来说明。[7]由于可以用分析的方法来解决方程，我们就无法快速看清非线性动态系统可以做什么。看清系统如何运作的唯一办法就是让它运作。在电脑上让系统运作非常简单（例如，银行账户的案例就可以用一个简单的电子数据表来解决），但如果用手工来做就会变得繁重异常。因此，在庞加莱的发现之后，科学家除了选择忽视非线性系统或用线性方程进行近似替代之外别无他法。对于非线性系统的研究被冷落了70余年，直到20世纪60年代至70年代才有起色，因为当时出现了新数学工具和电脑，研究才得以重启，非线性系统如今已成为物理学家的基本话题。[8]

这种发展非常重要，因为非线性系统在自然界中十分普遍，涵盖了从飞机机翼上方的湍流到气候、激光以及大脑中的突触刺激等诸多现象。非线性系统是如此之普遍，而真正的线性系统则相对罕见，这让数学家伊恩·斯图尔特（Ian Stewart）认为，物理学家专门将一个领域定义为“非线性系统”是非常愚蠢的，这就好比动物学家专门定义了一门“非大象动物学”——关于除大象之外的所有动物的研究。[9]

经济是复杂的，但不混沌

我们已经指出经济是一个动态系统，现在，我们又能进一步认定它是一个非线性系统。自杜尔哥在18世纪提出收益递减法则以来，经济学家已经承认经济中存在非线性关系。我们还能在生活中直观地看到非线性现象——从技术变革的加速到我们第5次看到同一则广告时注意力的降低。相关统计已经证明，微观经济数据中的非线性证据涵盖了从工业产值到失业人数等的种种现象。[10]尽管经济学早已承认非线性的存在，但曾经有一段时间，将非线性融入动态系统遇到了很大的困难。由于非线性动态系统之前并未获得充分理解，传统经济学家在历史上要么用静态系统中的非线性关系，要么用动态系统中的线性关系，这样方程才能被解开。

复杂经济学承认经济既是非线性的又是动态的，而且经济学家可以使用越来越先进的数学工具和电脑模拟对其进行分析。

关于“经济是否是混沌的”问题在20世纪80年代引起了人们的极大兴趣，尤其是股票市场是否混沌的问题（以及这个问题的答案是否能让所有人赚到许多钱）。[11]随后，人们针对这个问题开展了大量工作，当前的共识为“或许不是”。这个问题的答案很难确定，因为要想将混沌与其他复杂而非混沌的行为以及某些类型的随机行为区别开来并非易事。对于混沌的测试也需要大量的数据，金融数据也不足以获得确凿的答案。[12]大多数研究者认为，将经济称为混沌是一种过于简单而狭隘的归类。真正的混沌系统往往只有相对较少的变量及低程度的自由。将经济称为复杂系统更合适，正如我们之前指出的那样，经济领域拥有大量存量，并且流量在正反馈和负反馈的复杂网络中动态地连接着，这些反馈关系在不同的时间尺度里延迟和运作。这样的系统很有可能拥有惊人的自由度，能够收集非大象动物动物园里的全部行为。因此，我们可以说在某些时候，在某些维度中，经济或许会展示出混沌行为，但经济同样也会展示出增长、衰退、周期性极限环、准周期性极限环及大量其他行为。

如果我们将经济视为科学家所谓的多体问题（n-body problem），那么经济的全部动态复杂性就会变得十分清晰。庞加莱最初研究混沌的动机在于，1887年瑞典国王奥斯卡二世设立了一项奖金为2 500克朗的奖项，以寻求太阳系统是否稳定，或者说行星是否会撞向太阳或脱离太阳系飞向太空的答案。[13]这是一个极其难以解答的问题，不仅是因为太阳系跟围绕它运转的行星一样是动态的，也不仅因为重力是非线性的，而且还因为所有的相互作用是在10颗行星和太阳之间产生的。庞加莱意识到，他甚至连涉及三个主体的问题都无法解答，更遑论当时人们所知

的8颗行星和太阳。奥斯卡国王提出的问题直到20世纪才得以解决（你不用担心这个问题：在地球轨道变得不稳定之前，太阳就会爆炸），而多体问题依然无从下手。经济是一种大规模的多体问题，经济里的每一个个体都拥有自己的一套存量（储蓄、负债、技能等）和流量（收入、开支、学习等），就好比在高度简化的糖域中，每一位行为主体都拥有自己的存量（糖）和流量（糖消耗和消化）。正如糖域的动态是想象出来的经济行为主体之间互动的结果，现实经济的动态是数十亿人非线性互动的结果。经济系统的复杂性使得庞加莱的分类看起来太简单了——这不仅是一个多体问题，而且是70多亿体问题。

这样一来，大家就能理解为何经济预测是一项艰巨的工作了，其口碑甚至不如天气预报。对于初始状况的敏感性、路径依赖性以及高度动态复杂性使得经济像天气一样，只能进行短期预测而无法进行长期预测。近年来，通过卫星和雷达获得的更优数据以及通过电脑模型获得的更精细分析，天气预报员的业绩记录获得了显著提高。我们有理由期待，以后的经济预测也很有可能会随着更优的数据和模型的出现而同样获得提升。然而，正如天气的长期高度精准预测永远不可能实现一样，经济的长期高度精准预测也不可能实现。

经济长期预测的不可行性并不是该领域作为一门科学的发展障碍。如前所述，科学关乎解释，而非预测，并且对经济动态特性的理解将极大地帮助发展对于经济现象的可测试解释。

看不见的手偶尔会颤抖

来自麻省理工学院斯隆管理学院的约翰·斯特曼（John Sterman）教授花费了职业生涯的大部分时间，利用非线性动态技术来发展关于经济和商业现象的解释。[14]有一个问题引起了他的特别兴趣：为何许多商品会经历繁荣与萧条的循环？经历周期的行业类型囊括了全部范畴，用斯特曼的话说，“从飞机制造到锌处理”都包括在内（见图5-6）。[15]这些不同行业的共同之处在于，它们的价格和产业能力的周期震荡比基础需求或经济整体的震荡要激烈得多，因此，即使没有重大的缘由也能产生显著的效果。

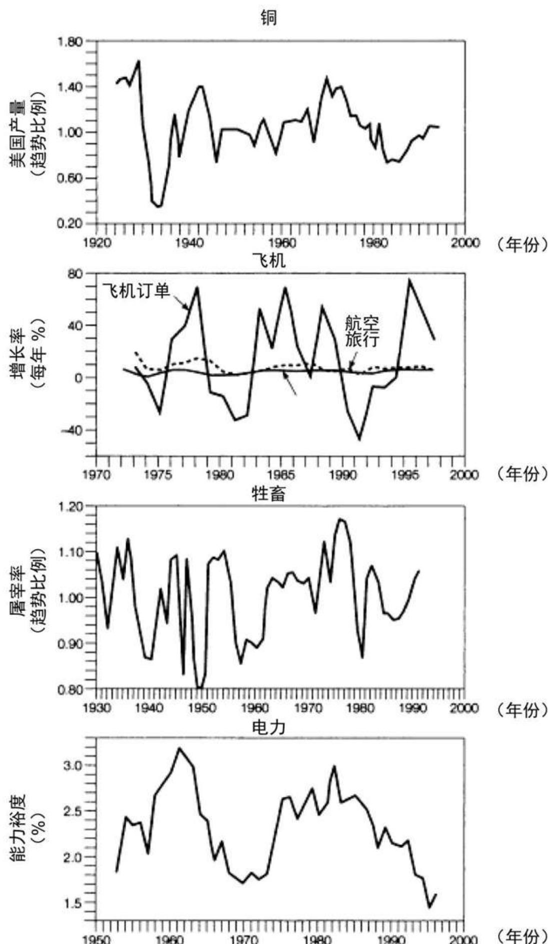


图5-6　周期性行业案例

资料来源：斯特曼（2000年）。铜和牲畜的数据来自美国商品研究局发布的《商品年鉴》，飞机数据来自皮尤-罗伯特联合公司、剑桥、MA，电力数据来自爱迪生电力研究所美国部分的统计摘要。

这些循环同样拥有一种有趣的特性，既不是特别规律也不是特别随机。通过观察它们，我们会看到，数据显然不仅是在随机波动，它们拥有有限的周期。但这些循环又不是十分有规律，它们的周期性并不完美。我们可以肯定地说，这些循环是复杂的。从商品贸易商到行业高管，许多人都尝试过对这些循环进行预测，但鲜有成功。斯特曼决定打造一个模型来调查是什么驱动着这种既不非常规律也不特别随机的行为。

传统经济学中最重要的负反馈是供求关系中的价格。假如需求增加，价格就会上涨，从而导致供应增加，这又会导致价格下跌直至供求

平衡，就像一台恒温器。然而，正如我们之前讨论的那样，传统经济学通常会假设所有事情都会同时发生，而忽略了时间延迟的作用。[16]因此，尽管供求平衡存在于经济学教科书中，现实世界里却充满着库存、产能过剩以及缓冲不平衡的其他存量。斯特曼假设各种缓冲存量调整速度的差异或许是商品动态循环背后的终极真相。

斯特曼决定建立一个简单商品市场的电脑模型，来看看是否可以复制这些循环的数据特征，从而验证自己的假设。[17]与传统经济学模型不同的是，斯特曼的模型（一种“系统动态”模拟）拥有明确的库存和产能存量、积极和负反馈循环、时间延迟和非线性关系。为了弄懂斯特曼的模型是如何工作的，我们可以假设你在一家普通制造企业负责生产小部件（存在于许多经济学入门课程里的一种神秘产品），企业的供应和需求的不平衡中存在着三种关键的缓冲库存。

首先，小部件的库存。库存充当着客户的不确定需求与工厂的生产之间的缓冲器。如果客户的订单量低于产量，你的库存就会增加；如果订单量高于产量，库存就会减少。库存几乎是实时调整的（虽然在接到订单和发货之间会有一点点延迟）。

其次，即刻可用的产能存量。如果在开始时小部件库存很少，你可以要求工厂加大生产。正常情况下，大多数工厂的产能利用率不到100%（通常为80%），这为工厂带来了一些短期的弹性产能。如果你打电话要求增加产能，工厂经理可以增加小部件生产线的工作时长，或增加工人的班次，或启用一条闲置的生产线，或将滞销小部件生产线转变成热销小部件生产线。提升短期产能的时间延迟可能是几小时，也可能是几个月，关键在于不能即时发生。产能回路的调整速度要慢于库存回路的调整速度。

最后，长期产能的总量。一旦所有生产线以最高速度运转，并且利用率达到了100%，那么唯一能够增加产出的办法就是建造新的生产线或新的工厂。增加新的长期产能的速度慢于提高已有产能的利用率的速度。建造新工厂、雇用更多员工等往往需要几个月甚至几年的时间。

斯特曼的模型包含了三个调整速度各不相同的反馈环结构。要想弄清楚这个结构是如何随着时间变化而发挥作用的，你可以想象自己是小部件生产线的经理，每周都会收到一份生产报告。当你坐下开始阅读本周的报告时，你注意到小部件的需求增加了，库存下降了一点。库存会随着需求的随机波动而出现变化，所以你要考虑的第一个问题就是是否要求工厂提高产量。这时往往会出现一个自然的反应，那就是在付出昂

贵代价增加产量前（有可能造成产量过剩）再等一等，看看上涨的需求是否会持续下去，或者仅是昙花一现。随后会出现第二个问题：是否应该涨价？这时你希望弄清需求的增长是否真切。涨价需要时间和工作量（新的价格需要在电脑上进行运算，客户合同需要变更，等等），并且永远伴随着客户流失的风险。因此，你会决定再多等一段时间。

在下一份报告中，你看到需求持续增加，至此你开始相信确实存在这样一个趋势。现在库存已经下降到一个危险的低点，你不得不针对部分客户采取延期交货的措施。你开始采取行动，并提出增加产量的要求。集团总部需要突破官僚制度，工厂需要一些时间来进行调整，真正提高产量要到几周之后。与此同时，需求仍在持续增加，但由于客户用掉了库存，于是发生了缺货和积压订单。考虑到旺盛的需求和隐约出现的产能不足（你还注意到竞争对手也正在经历相同的状况），你决定是时候涨价了。

价格的上涨不会立刻对需求产生影响。由于货物短缺，客户的库存已经耗尽，他们无法马上找到替代品或其他资源，因此他们愿意（至少在短期内愿意）支付更高的价格。有了不断增加的需求和更高的价格，你的损益表看起来漂亮极了，你成了首席执行官眼中的英雄。你唯一面临的挑战在于必须生产出足够的产品。工厂的运行效率接近100%，与此同时，你的订单和市场份额流失给了拥有更高产能的竞争对手。于是，你会向管理委员会提出投资建立新厂的提议。鉴于你在业务上的英雄表现，他们会批准你的提议。新厂址破土动工，开始建厂。

在距离新厂建成投产还有半年左右的时候，你注意到一个令人担忧的趋势，需求的增长趋势开始放缓，数个月的价格上涨和产品短缺开始产生影响，客户找到了你们产品的替代品以及减少用量的办法。此外，在你加紧生产以满足需求的时候，你的竞争对手们也在这么做，此时便会出现很多库存。由于需求的放缓以及竞争对手增加了现货，你从销售队伍那里听到了降价的呼吁，再次感受到了压力。你同意了销售队伍的建议，批准打折。

到了新工厂举办开业典礼的那一天，正当首席执行官剪彩完毕，你就收到了一份让你的热血变凉的报告。需求在进一步放缓，更糟糕的是，价格正在发生自由落体式下滑。行业里的所有人都有一大堆库存，都在大幅度打折以便维持运转。尽管面临着价格下跌和越来越难看的损益表，你还得继续前进，激发新工厂的产能。毕竟，工厂已经建成——大部分都是沉没成本，也没有什么别的用途。即便产品价格很低，至少

还能覆盖用工成本和材料成本。

几个月之后，你参加了一次行业展销会，与其他高管坐在吧台上，一起哀叹行业的不景气。结果发现，在景气的时候，他们的所作所为跟你一样，现在他们也有崭新的工厂即将投产。你曾怀疑过竞争对手们在建新的工厂，但由于行业保密性，当时你并不知道到底有几座工厂在筹备中，但你认为不断增加的需求可以消化这些增长，你还假定人们不会疯狂到过度投资建厂。

到了第二年，行业产能泛滥，价格崩溃，整个循环开始往反方向运转。一段时间之后，随着亏损不断增加，行业开始关闭生产线，减少班次，从而降低产能。随着“流血事件”的进一步恶化，同行业的公司都开始采取下一步行动，即关闭整个工厂，解雇大部分员工。通常情况下，这需要数年时间才能将过剩的产能从系统中挤出去，但这已不再需要你的担忧，因为你已经在某一轮的裁员中被解雇了。

几年之后，一位年轻的经理坐到了你昔日的位子上。在上一个周期的时候，她还只是一个大学生，根本不记得有过这回事。有一天，她坐在座位上，收到了最新的小部件生产报告，她很激动地看到需求出现了良好的增长势头……

这个动态故事就是斯特曼模型的本质。在运行模型的时候，斯特曼发现模型会产生商品周期，从统计上说，这与现实世界中的周期有着许多重要的相似之处。[18]模型表明，反馈环的不同时间尺度以及人为误差会使得这样的周期几乎不可避免。斯特曼及其同事马克·帕奇（Mark Paich）还进行了另外一项实验：让真人扮演经理的角色并尝试使周期最小化。[19]然而，研究者发现，在处理复杂反馈和不同时长延迟的系统时，人们在心理上会经历一段难熬的时间，会遭遇一段非常困难的时间（就像反应很慢的淋浴喷头一样）。斯特曼的模型和实验给我们的一个启示在于，缩短周期的唯一办法是改变系统本身的结构。比如，我们可以减少系统的时间延迟，例如增加新产能的时间；降低产能的沉重程度，例如建造“迷你工坊”而非大型工厂；对客户订单获得更多前瞻性预见，或者增加行业实际已有的以及正在建设的产能的透明度。

斯特曼的工作有意思的一个方面是，它强调对我们的心智模型来说，要想理解和发展关于非线性动态系统的直觉非常困难。[20]问题不在

于人类的愚笨，而在于我们的大脑没有被设计成按照这种方式进行思考。经理和政策制定者已经尽其所能，但他们的行为通常会产生无法预测的结果，甚至会让情况更糟。在下一章，我们将较为一般性地讨论人类的认知缺陷。在后续章节中，我们将进一步研究协调人类行为和经济的动态结构如何有助于解释复杂现象，这些现象包括从经济系统的增长、衰退循环到股票市场的易变性。

06

视角二：行为主体是通过归纳推理进行决策的

任何经济学理论的核心中都有关于人类行为的理论。经济最终是为人服务的，有关经济行为的理论必须回答的问题包括：我们是如何进行经济决策的？我们运用的是什么类型的信息？有哪些我们更擅长或更不擅长的决策类型？

人类行为可以通过科学的办法加以理解，许多人对这一观点持有天然的排斥态度。毕竟，我们怎么可能理解人脑这种复杂、不可预测，并且还具有情绪、创造力、想象力的东西呢？我们如何能够将可以爱、恨、创作诗歌的人类放到一个数学公式里呢？在艾萨克·阿西莫夫（Isaac Asimov）的经典科幻小说里，身处遥远未来的科学家发明了一个名为心理历史学的领域，在这个领域，他们能够将个人的行为简化为数学方程，从而使得这些科学家可以对历史的未来进程进行准确的预测。心理历史学听起来像是经济学家的终极幻想。然而，我们并不需要心理历史学之类的人类行为预测理论来理解经济的运转原理。我们的标准要低得多，所需要的不过是找到人类行为的一些基本规律，然后通过理解这些微观层面的规律更好地理解宏观层面的经济行为。我们将在本章看到，人类的行为中充满着规律。

正如之前讨论的那样，所有科学理论都是现实世界的近似和简化，因此，我们创建关于人类行为的经济模型，其目标并不在于创造一个完美的复制品，它更像是一张优秀的地图，如实反映地貌的基本特征，但剔除了无关的细节。此外，模型不应该加入任何不存在于现实世界里的东西（例如，模型不应该赋予我们看透他人心思或者其他真人所不具备的能力），也不应该与我们所知的真相矛盾。

糖域就为我们展现了一个这样的简化案例。糖域中行为主体的行为比真人的行为要简单得多。这些行为主体只会进食、移动、繁殖和交易，甚至于他们做这些事情的方式几乎简单到不值一提。然而，物理移动、繁殖和交易全都是真实的人类活动，并且都有可能与理解经济现象有关。糖域中不存在任何真人不会或不能做的事情。

经济学家喜欢将他们的人类行为标准模型称为“经济人”（Homo economicus，即economic　man）。在本章中，我们将会看到，相对于传统的经济手段，经济人的复杂经济手段发生了彻底的变化。正如我们之前简要叙述的那样，人类行为的传统模型是传统经济学理论中比较具有争议性的假设之一。现在我们将进一步对此进行检视，并且我将论证传统经济学的经济人是一张“糟糕的地图”：它既忽视了真实人类行为中的重要细节，又增添了真实人类身上并不存在的关键特征。

斯波克去购物

《星际迷航》里的瓦肯星人斯波克就是一个完美的理性人形象。斯波克能够记住圆周率的小数点后50位，可以进行无比复杂的计算，而且这一切都是在敌人相位武器的攻击下进行的，整个过程没有丝毫的情绪波动。斯波克就是传统经济学刻画人类行为的一个很好案例。假设你走进一家当地的杂货店，看见一些西红柿，接下来这些段落描述的将是传统经济学认为你在决定是否购买西红柿的过程中将会产生的所思所想。[1]

相对于世界上其他一切你都有可能购买的东西，例如面包、牛奶和西班牙之旅，你对购买西红柿情有独钟。此外，你对于在未来某个时间点可能购买的所有东西都有着明确的偏好，由于未来是不确定的，所以你为这些潜在的购买需求赋予了概率。比如，你认为在未来两年内，厨房里的架子散掉的可能性为23%，到时你需要花1.2美元购买螺丝来修好它。1.2美元的折现值为1美元，再乘以23%的概率，相当于未来可能需要修理的预期值为23美分。而今天，你必须对想要购买的西红柿，以及余生中有可能要购买的其他所有东西进行权衡。在传统经济学的模型中，所有这些明确的偏好都有着非常合乎逻辑的排序。因此，如果你喜欢西红柿胜过胡萝卜，喜欢胡萝卜胜过青豆，你一定会先买西红柿、后买青豆。同样，如果你喜欢西红柿胜过胡萝卜，也不会因为看见了青豆而突然购买胡萝卜。

传统经济学还假设你对购买西红柿的花费有着明确的预算。为了得出这个预算，你必须对余生的收入有深思熟虑的预期，并且基于此对当前的预算进行最优化计算。换而言之，你有可能意识到你花在西红柿上的钱可能在退休后更有价值，从而放弃购买。当然，这假定了你未来的收入将被用来投资金融资产的绝佳对冲投资组合，并且它还考虑到了你将会在65岁退休的可能性，以及你对未来利息、通胀和日元兑美元汇率

的预期进行了精确计算。当你站在那儿，看着那些色泽诱人的西红柿，你会将所有这些信息装进脑海里，展开一次狡猾的、无比复杂的最优化计算，来权衡所有这些因素，最终得到一个绝对最优的答案——购买还是放弃！跟斯波克有异曲同工之处的方法被称为完全理性模型，此乃传统经济学的根本假设之一。[2]

相反，新兴复杂经济学关于行为的核心观点里存在着另外一种被称为归纳理性的方法，它认为情况会是这样："嗯……今天的西红柿看起来真好，真新鲜。今晚我有点儿想吃沙拉……价格还可以。"接着你就会将西红柿装进购物车里……

认知失调

像传统经济学的许多其他理论一样，完全理性也要回溯到我们的老朋友瓦尔拉斯和杰文斯那里。他们并没有对人类行为进行实际研究就得出了这个模型。没有人见过有人会做这样的计算，或作出这样的行为。将完全理性作为一种假设是为了将经济学装进19世纪的均衡框架里。

正如我们之前讨论过的那样，在均衡经济学的球与碗框架中，完全理性发挥着类似于重力的作用。如果我们知道人们行为的约束条件，并且假设每个人的行为都是完全理性的，那么他们将会对这些约束条件作出一致的反应。人们的决定因此会变得可预测，然后使系统实现均衡。非常值得一提的是，并不是有人认为传统经济学的关键行为假设很好地描述了实际的人类行为，因此它们才会被开发出来，它们是被用来使均衡框架中的数学发挥作用的。瓦尔拉斯以及后来的经济学家试图在事后为假设中缺乏现实主义找到合理的解释，他们分辩称，即便完全理性的假设不能很好地描述人类的实际行为（用经济学的术语来说不是"实证模型"），它也可以被看作阐述应该如何行动的描述（即"规范"描述）。我们可以构建一个完全理性的经济学模型，然后看看现实距离这个模型有多远。

如果说完全理性描述的是人们应该如何行动，那么仍然存在两个问题。首先，如果其他所有人都按照不一样的方式行事，那么剩余的一个人就成了唯一"完全理性"的人，这样的说法未必成立。我们在讨论动力学的时候曾说过，时机就是一切。举个例子来说，如果某人在1998年抛售了跌价的互联网股票，但他要到2000年才会知道这种做法是正确的，在此之前他还要经历两年孤独的亏钱时光。其次，我们很快将会看到，

完全理性不是一个好的规范模型，因为即便人们希望如此行事，事实上他们也做不到。

正如第3章中讨论的那样，完全理性的假设在引入之初就受到了挑战和批评，但经济学家仍在继续使用这些假设，这是因为完全理性为模型增添了数学因素，并且没有人有更好的方案。20世纪50年代，赫伯特·西蒙及其在卡耐基梅隆大学的同事詹姆斯·马奇（James March）和理查德·西尔特（Richard Cyert）提出了一个更加直接的挑战。他们做了一些颇具颠覆性的事情——真实地观察和研究人们的决策过程，尤其是真实公司里真实经理的决策。这些经历，再加上心理学家对他们工作的解读，使得他们相信真实世界的行为与经济学教科书中所呈现出来的行为没有半分相似之处。西蒙提出了一个关于决策的博弈论——有限理性（bounded rationality）。[3]大体上，这个理论认为人类在经济上是利己且聪明的，但没有聪明到完全理性的程度。西蒙的理论考虑到了人们缺乏完全的信息，人类大脑的信息处理能力虽然强大但是有限。如前所述，西蒙提出人们不会遵循完全理性，而会满足于利用已知的信息尽量做到最好。

因为这些工作，西蒙赢得了1978年的诺贝尔经济学奖，并且绝大部分经济学家如今提到他时仍然充满敬意。尽管如此，他的成果却对于传统经济学核心力量的影响甚小。[4]大部分标准经济学教科书中甚至不会提到西蒙的有限理性。[5]之所以存在这种忽略，其中一个重要的原因是，经过多年各式各样的尝试，研究者依然未能将西蒙的理念转化为数学模型。[6]

继西蒙的开创性工作之后，两代经济学家和心理学家开始收集大量经验证据和实验证据，用以对抗传统经济学的经济人模型。20世纪70年代，普林斯顿大学的丹尼尔·卡尼曼（Daniel Kahneman）与斯坦福大学的阿莫斯·特沃斯基发表了一系列奠基性论文，用以证明真实的人类决策是如何与传统经济学模型中某些最基本的假设存在直接冲突的。[7]20世纪80年代和90年代，诸多研究者，包括保罗·斯洛维奇（Paul Slovic）、科林·卡默勒（Colin Camerer）、阿尔文·罗思（Alvin Roth）、莱因哈德·萨尔腾（Reinhard Selten）、理查德·塞勒（Richard Thaler）、乔治·洛温斯坦（George Lowenstein）、马修·拉宾（Mathew Rabin）、德拉赞·普莱克（Drazen Prelec）、赫伯特·金蒂斯（Herbert Gintis）、恩斯特·费尔（Ernst Fehr）、约翰·卡格尔（John Kagel）和弗农·史密斯，共同创建了今天大家所熟知的行为经济学。[8]这类研究一次又一次地确认了真

实人类不会像传统理论中所描述的那样形成偏好、判断风险或进行决策。新千年后，行为经济学进入主流视野。2002年，丹尼尔·卡尼曼与弗农·史密斯获得诺贝尔经济学奖（由于诺贝尔奖不能授予已故之人，所以许多人认为卡尼曼的长期合作者、已故的阿莫斯·特沃斯基是被遗漏的“第三人”）。[9]

行为经济学的兴起使得该领域进入了一个认知失调的怪异状态：许多经济学家承认对完美理性批评的合理性，但他们还在坚持不懈地使用传统的假设，因为他们缺少替代方案，缺少一个被正式提出的模型。后来，随着行为经济学家、心理学家和计算机科学家之间的合作不断加强，另外一种模型开始浮现。在接下来的几节里，我们将会看到由行为经济学家找出的部分证据，然后梳理出一个新经济人的模样。

你这头自私的猪！

假设你乘飞机旅行，座位挨着过道，旁边是一位看上去有点古怪的女士，她坐在中间的座位上。[10]那位女士旁边是一个靠窗的座位，上面坐着一位你不认识的商务人士。在飞到一半的时候，那位看起来古怪的女士对你们俩说，她是个有钱人，长途飞行让她感到十分无聊。接着她提出，她会给你和那位商务人士5 000美元，前提是你们能够达成一致的分配方案。过程可能会是这样的：由商务人士决定怎么分钱，他会分给你一部分。如果你接受他的提议，那么你们将各自得到各自的部分。如果你拒绝，那么你们都得不到钱。商务人士想了一想，继而转向你说：“我的提议是我拿4 990美元，你拿10美元。”你会怎么做呢？是接受还是拒绝？

如果你跟其他大多数人一样，那么你会拒绝这个提议。为什么？因为它不公平——商务人士就是一头自私的猪！然而从经济学的角度来说，拒绝这个提议完全是非理性的。无论对方提出什么样的方案，哪怕是给你1美元，你都应该接受。毕竟，即便自私的商人获得了4 999美元，你得到1美元也比什么都没得到要好。

类似这样的最后通牒博弈已经在各个地区的真实人群中（例如日本、以色列、斯洛文尼亚、智利、津巴布韦、印度尼西亚和美国）发生了数千次。这种博弈还在大学生、商人以及坦桑尼亚的牧民中间发生过。结果惊人地一致，人们的实际行为与经济学教科书中的描述大相径庭，大多数人都会拒绝接受他们认为不公平的提议，哪怕这意味着他们

要放弃近在咫尺的钱财（不同文化中的人对不公平的定义是有差异的，但通常是指少于30%）。[11]那么，这是否能说明人们是情绪化的、非理性的呢？

某些经济学家会提出，拒绝10美元依然是完全理性的。我们可以再来看一下一些经济学家提出的所谓的“重复博弈”。通过拒绝提议来惩罚那位商人，你或许在不知不觉地假设你们在未来还会有接触，并且既然商人知道了你不是容易受骗的傻瓜，那么他就更有可能在未来的交易中给你分配更大的份额。因此，你现在理性地拒绝10美元，是在为了以后获得更多的收益做打算。然而，这种解释经不起推敲。在一系列经过谨慎设计的实验中，来自苏黎世大学的恩斯特·费尔与西蒙·盖希特（Simon Gachter）证明，即便参与者知道未来双方不再有交集，并且现在的损失不会换来未来的收益，还是会出现同样的行为。[12]

人们的公平意识和互惠意识会让他们产生一种惩罚对他们不公之人的渴望。此外，我们还会自然而然地回报那些帮助我们、为我们付出的人。比如，研究就餐行为的研究者发现，如果服务生在给账单的时候留下一枚糖果，那么他获得的小费要比别人高18%（留下两枚糖果得到的小费更高）。[13]

传统经济学模型假设人们只关心经济决策的结果，而会忽略决策的过程，而像议价、公平和胁迫这样的事情根本不存在于模型之中。此外，模型还假设人们只关心个人的得失，丝毫不理会他人。最后通牒博弈实验提供了反对这些假设的直接证据。在近期有关行为研究的一项调查中，来自马萨诸塞大学和圣塔菲研究所的赫伯特·金蒂斯以及一组研究者指出，亚当·斯密在其著作《国富论》中将人类刻画成了自私、金钱至上的生物。[14]金蒂斯及其同事评论说，许多人忘记亚当·斯密还写过另外一本书——《道德情操论》，在此书中，斯密更加细致入微地展示了既自私又慷慨的人性。最后通牒博弈和其他实验所得到的证据表明，亚当·斯密的后一种观点是正确的。对于公平和互惠，人类具有根深蒂固的规则意识，它们远远超越了计算出来的“理性”。金蒂斯和他的同事们注意到，人类是“有条件的合作者”，只要对方慷慨，自己也会慷慨，同时人类也是“利他的惩罚者”，人们会回击那些作出不公正行为的人，甚至不惜以自己的当前利益作为代价。

我们既是有条件的合作者，又是利他的惩罚者，这不应当令人感到惊奇。原始人类在长达200万年的时间里，都生活在合作行为与生存高度相关的小群体里。如今，人类依然生活在社会互动的网络中，在这样

的环境中选择互惠非常重要——如果你帮我梳毛，我也帮你梳毛。如果能够互相帮助，那么我们的处境都会变得更好，但这也会让那些光占便宜不付出的人钻了空子。少量“搭便车”的行为可以容忍，但如果泛滥开来，互相梳毛的系统就会崩溃，这会让每个人的处境都每况愈下，这也解释了为什么我们会拥有根深蒂固的奖励合作、惩罚投机的行为。[15]部分经济学家或许会将这些行为看作非理性行为，我们将在后续的章节中看到，这些行为实际上为社会合作提供了基础，而社会合作是人类创造财富的关键。

人非圣贤，孰能无过

真实人类与传统经济学模型的另一个重要区别在于人类会犯错误。传统模型假定人类是像斯波克一样的行为主体，是完美的，从不犯错。假使他们会犯错，这些错误也不过是随机分布在正确答案周围的噪音。换言之，传统经济学模型中的人类不会犯相同的错误或偏差。

行为经济学的观点偏向于“人非圣贤，孰能无过”。哥伦比亚大学的行为研究专家乔恩·埃尔斯特（Jon Elster）甚至声称：“最有可能进行无偏差认知评估的是抑郁症患者。”[16]研究者发现，正常的、非抑郁状态的人们身上所具有的普遍错误和偏差包括：[17]

- 框架偏差。问题的设计方式将会影响我们对问题的看法。比如，请比较以下两个问题：“英国是否应该使用欧元？”“英国是否应该废止英镑？”在完全理性的情况下，这种设计应该不会造成影响。
- 以偏概全。人类有一个坏习惯，即从很小的、具有偏差的样本里得出普遍性结论。比如，你跟办公室里的三位同事进行交谈，而他们碰巧在那一天过得都不顺心，于是你可能就会得出公司即将破产的结论。
- 可用性偏差。人们倾向于利用容易获取的数据来做决策，而不是寻找那些有助于作出良好决策的数据。也就是说，人们“在路灯下寻找丢失的钥匙”是因为那里的灯光最亮。
- 风险判断困难。大多数人在论证可能性和评估风险的时候都会遇到困难。2000年10月，一列火车在英格兰哈特菲尔德失事，造成4人死亡、34人受伤。作为回应，英国政府

提出要在铁路安全方面额外投入20亿美元。然而，正如《经济学人》指出的那样，相对于其他交通方式，人们在火车事故中丧生的实际可能性极低，而这样的支出意味着在每位预期可挽救的铁路旅客身上的花费是道路安全人均花费的150倍。[18]

- 迷信的推理。我们往往只能看到最直接的原因并因此混淆随机因素与因果联系。从体育明星穿着“幸运之袜”到政府通过增加解雇的难度来降低失业率，这样的案例比比皆是。

- 心理账户。传统经济学对所有的钱都一视同仁。然而，人们通常会把钱放进不同的心理隔间里。比如，许多人会每月将一部分钱划拨到退休账户上，哪怕现在他们的信用卡透支已经非常严重。从经济方面来说，这种行为并不理性，因为退休账户上的投资回报（即便节省了税费）很有可能抵不上信用卡利息。尽管如此，人们还是会将他们的退休缴款视作神圣不可侵犯的，将它与眼前的开销分隔开来。对不同类型的开销进行心理区分由来已久，有位人类学家甚至在非洲的卢奥部落中也发现了这种行为。[19]

无法计算

面对人类这些意志薄弱的行为，传统经济学家有时候会回应说，诚然，大多数人都会偏离理性，但市场需要的不过是一小部分、甚至是一位超级理性的参与者，他能利用其他所有人的错误，将市场推动到完全理性均衡的状态。即便大多数投资人都是傻瓜，只有少数几位像斯波克一样的交易员会利用其他所有人的愚笨套利，赚得盆满钵满，使价格趋于完全理性所预测的程度。[20]抛开实现套利的实际性不说（是否有人拥有足够的信息和资本这么做？它是否能在不那么“完全”的市场，例如伦敦番茄酱市场中奏效？时间迟滞、调查成本等又该怎么办呢？），另外还有一个大问题：是否有人（包括斯波克先生）聪明到能实现这一点？

1985年，数学家阿兰·刘易斯（Alain Lewis）利用计算理论中的一些复杂技巧证明，没有人（哪怕是最聪明的套利者）可以进行完全理性理论所描述的计算。[21]刘易斯证明了经济学理论中完全理性或许可能存在，但在任何实际场景中都不可能存在，因为它是不可计算的。计算理

论中有一个概念叫作图灵机，以数学家阿兰·图灵的名字命名。[22]如果某事物可以在图灵机上进行计算，那么至少在理论上是有可能制造物理计算机对此进行计算的。然而，如果某事物不能通过图灵机进行计算，那么就没有一台计算机能够解决这个问题，无论它多强大，哪怕它的体积跟宇宙一样大。刘易斯证明了传统经济学家所定义的完全理性不能通过图灵机进行计算。

尽管认知科学家和哲学家还在辩论从技术上说我们的大脑是否是图灵机，但刘易斯的研究表明，完全理性的计算需求是巨大的（这与之前讨论中提到的经济需要经过百万兆年才能达到均衡的情况相类似）[23]，这种量级和类型的计算能力是我们的大脑不可能达到的。

阿瑟的酒吧问题

除了无法计算，许多类型的经济问题都不存在完全理性的解决方案——理论上和实际上都没有。[24]起初，布莱恩·阿瑟也是圣塔菲研究所会议的参与者之一，他也曾担任过圣塔菲研究所经济学项目的主管。阿瑟通过以下案例解释了以上观点。[25]圣塔菲研究所有一个很受欢迎的酒吧——爱尔法鲁尔酒吧，每周四晚上那儿都会现场演奏爱尔兰音乐。这个酒吧不大，如果酒吧里的人数不超过60人，那么你就能度过一个舒适而又愉快的夜晚。而如果人数超过60人，那里就会过于拥挤而令人感到不舒服。如果你预计人数小于等于60人，你会决定在周四晚上前往酒吧；如果你认为人数将超过60人，你会选择留在家里。你无法与其他可能前往酒吧的人取得联系，并且你也不能打电话问埃尔·法罗尔酒吧到场人数有多少。我们假设所有人都按照这个方式来决定去或不去酒吧，那么你会去酒吧还是留在家里？你会如何进行抉择呢？

结果发现，人们对于这个问题根本没有完全理性的解决方案。这其中包含着无限循环——A的行动取决于A对B行动的预期，而B的行动又反过来取决于B对A行动的预期，如此循环往复。由于人们无法通过分析得到答案，针对埃尔·法罗尔酒吧问题进行决策的唯一办法就是让参与者们回顾过去去酒吧的经历，看看能否找到某种模型，然后进行判断：“过去两个周四我都去了，人不算太多，所以我会再去。”

阿瑟用计算机模拟遵循经验法则的行为主体，发现酒吧永远不会达到均衡。在某些夜晚，酒吧很满，有时是半空，但酒吧人数波动的平均值为60（见图6-1）。传统经济学家或许会对着这个图说：“啊哈，它确

实达到了完全理性的到场率，也就是达到了酒吧舒适范围内的最大容纳度60人，因此我们可以说它找到了一种均衡，虽然周围有一些噪音。”然而，这种说法并不正确。平均60人到场并不是一种均衡，并且波动也不能仅仅被视为杂音而被排除。在该系统中，除了初始条件，没有什么是随机的，因此这个系统完全是确定的、零噪音的。正确的解读应该是：这个问题没有均衡，没有定点吸引子。[26]参与者在利用各种各样的经验法则，这种变动创造了一个接近60的动态吸引子，并且到场人数的高度波动性是由行为主体之间互动这个内因产生的，这种波动性永远不会停止——到场人数将永远会大幅度振动，并且永远不会收敛到均衡。

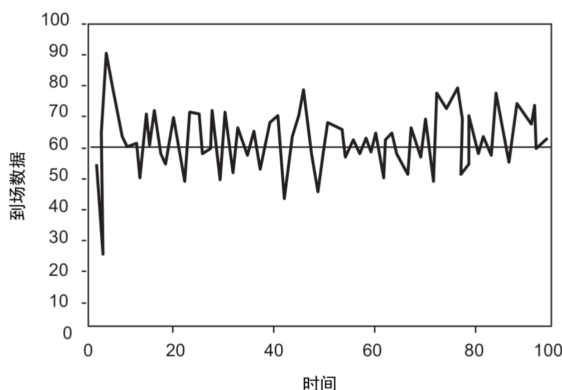


图6-1　埃尔·法罗尔酒吧的到场人数

资料来源：数据来自布莱恩·阿瑟（1994年）。

阿瑟指出，许多经济决策都是动态、自我指涉且定义不清晰的。A公司在考虑降价时不仅要担心自身的处境，还会顾及B公司和C公司将会如何行动，这两家公司反过来也会在某种程度上依据A公司的做法来制定战略。采用一种新的技术标准、在市场上为一款新产品寻找定位或者评估一支股票的价值，这些行为中都被加入了自我指涉的期待。人们勉强作出决策的唯一办法是从过去的经验中寻找模型和经验法则，这样一个由进化的、互动的、混乱的行为主体所组建起来的世界永远无法达到均衡。阿瑟的酒吧问题产生了一种有趣的可能性，即我们在现实世界的经济中看到的许多波动或许是由人们决策原则的动态性，而非外源性的随机振动引发的。[27]

归纳理性

这告诉我们一个什么道理呢？完全理性的标签是一种不幸的、充满情绪的措辞，因为它隐含着另外一种非完美和非理性的可能性，而没有人希望自己是不完美和不理性的，经济学家最不希望如此。那么我们唯一的选择是不是就只剩下举起双手作投降状，念叨人类变化无常的种种行径，然后放弃理解经济人的尝试呢？幸运的是，我们还可以用另外一种方式看待人类行为——一种能够为我们的新观点提供基础的方式，那就是现代认知科学。

认知科学是贴在人类大脑“软件”（相对于大脑“硬件”而言）研究领域的标签。该领域涉及多个方面，包括神经科学、心理学、人工智能、语言学、进化理论、人类学以及哲学。[28] 21世纪之初，这是科学界进步最快的领域。

现代认知科学是从两个角度看待人类大脑的。首先，大脑被认为是一个“信息处理器官”，一个“拥有计算能力的物体”。[29]然而，认知科学家不喜欢计算这个词，因为它拥有一种负面的隐含意义，隐约指向放在你桌上的那个笨拙的、白色（或黑色）的、有着塑料外壳的物体，而它唯一已知的情绪是幸灾乐祸的残酷——它会恰好在重要的截止日期之前死机。尽管我们的大脑也会计算，但它与人造计算机的计算方式截然不同。信息处理器官这个词隐含的意思是，为思想提供防护外壳的脑袋是生物设备的一部分，它拥有一项专门的功能：信息处理。这个词语还意味着，尽管大脑的复杂程度无与伦比，但它们并没有什么神奇的或不可知之处。事实上，人类的大脑是一个能被科学认知的实际物体，哪怕了解它需要花费很长时间。其次，现代认知科学是从进化角度理解大脑的。经济人的信息处理器官经过了自然选择力量的塑造，是人类物种的历史和环境的产物。我们的大脑不是由工程师在一张干净的白纸上设计出来的，而是由进化设计出来的——从灵长类动物数千万年的进化，再到原始人约200万年的进化。

认知科学研究发现，人类大脑所拥有的处理信息和学习的特长是无与伦比的，与完全理性所描绘的场景截然不同。比如，人类或许不太擅长计算长方程式，但他们是了不起的故事讲述者和聆听者。西北大学学习科学研究所主任、原耶鲁大学人工智能实验室主任罗杰·肖克（Roger Schank）开展了一系列围绕人类大脑的理解、记忆和交流过程的研究。[30]正如柏拉图所言，“会讲故事的人统治社会”。下次跟朋友共进晚餐的时候，你可以闲坐一会儿，观察一下现场的情况。大家都在做什么呢？大多数的夜晚，人们都在交换故事：有趣的故事、悲伤的故事、他人的

故事、报纸上看到的故事，等等。我们为什么会这么做？为什么讲故事和听故事对于人类的思考方式来说如此重要？

故事对于人类来说至关重要，这是因为我们处理信息的主要方式是归纳。归纳的本质是通过模型认知进行推理，以及根据证据中的多数情况进行总结。比如，尽管没有人看见管家行凶，但是他的指纹留在刀上，有人看见他离开案发现场，并且他有行凶的动机，因此，凶杀案肯定是管家干的。人们无法从逻辑上证明管家行凶，从逻辑来说凶手有可能是别人，毕竟没有人亲眼看见是管家拿刀杀人，但证据模型会让我们通过归纳得出结论：就是管家干的。

我们之所以喜欢故事，是因为故事就是我们的归纳思考机器的食粮，是我们寻找模型的材料。故事是人类学习的一种方式。举个例子来说，通过阅读莎士比亚的作品，我们可以学到关于爱情和家庭关系的各种有用的经验教训。比如你的父亲突然去世，而你的母亲跟你的叔叔结婚了，那么请你务必保持怀疑之心。畅销的商业图书通常是关于成功的个人或公司的故事，比如人人都想阅读关于杰克·韦尔奇（Jack Welch）[(9)]或比尔·盖茨如何“干了一票”的故事，希望收集一些关于成功的模式。

人类尤其擅长归纳模型认知的两个方面。首先是通过比喻和类比将新经验与旧模型联系起来。[31]下次你参加会议的时候，可以观察一下人们利用类别进行推理是多么频繁，他们会说“这跟1987年发生的行业洗牌是一样的”、“这位顾客让我想到了X公司”或类似的话语。20世纪90年代初，互联网第一次进入商业应用场景时，人们都在努力按照像或不像电视、收音机、杂志、软件以及电话对它进行定义。

其次，人类不仅是优秀的模式识别者，还是优秀的模式完善者。我们的大脑很擅长填补缺失信息所造成的裂缝。完善模式以及从高度不完整的信息中得出结论的能力使我们能在快速变化和模糊的环境中进行快速决策。

模式识别和讲故事与我们的认知密不可分，以至于人们会从完全随机的数据中寻找模式和构建故事。体育解说员和粉丝都热衷于讲述细节性的故事，如为何某人会连连取胜、进很多球，或者打出许多本垒打。在一个著名的分析案例中，托马斯·季洛维奇（Thomas Gilovich）、罗伯特·瓦罗内（Robert Vallone）与阿莫斯·特沃斯基观察了这种现象，证明绝大多数所谓的“热手效应”完全可以用随机因素来解释——如果在比赛中有足够多的运动员，其中某人在某个时间段注定会连续得分。人们是在通过编造故事来解释他们所认定的模式。[32]

为何"深蓝"系不了鞋带？

与归纳相对的是演绎。演绎是这样一个推理过程：结论必须从一系列前提出发，根据逻辑推断而来，比如，"苏格拉底是人，所有人都会死，所以苏格拉底也会死"。人类在使用归纳法的同时也在使用演绎法，但是用得不如归纳法熟练。相对于演绎法，人类在使用归纳法时往往技高一筹，计算机的情况恰恰相反。我们都能即时识别他人的面孔（这是一项归纳任务），但大部分人在进行以下演绎计算时会觉得极为困难：（239.46×0.48+6.03）÷120.9708。一个便携计算器就能够快速而准确地完成后一个任务（答案是1），然而，对于功能强大的计算机来说，准确识别面孔也是一项极其困难的编程挑战。

我们偶尔也会利用演绎作为强大归纳能力的补充，尤其在我们被难住、在模式仓库里找不到答案或者对于归纳直觉所给出的答案不够自信时。然后，我们就会低调而痛苦地（并非总是准确无误）利用逻辑推理来寻找解决问题的办法。在这个过程中，我们发明了铅笔和纸、代数、算盘、计算器、计算机，甚至是科学方法本身。一旦开始计算，经验就会进入模式仓库，这样我们就不用重复完整的过程。我们还不断地评估基于模式的判断成功与否，这样就可以从经验中学习。

象棋能够很好地展示人类的归纳与演绎思维是如何进行协作的。顶级的国际象棋棋手会利用归纳法来全局观察棋盘上的模式，同时他们还会利用演绎法来分析具体的局部方位。印第安纳大学认知科学项目主任安迪·克拉克（Andy Clark）指出，大师级的象棋棋手能在一瞥之间分辨出大约5万种不同的棋盘摆位，[33]他们的归纳模式仓库非常丰富，能够自动地对许多情况作出反应，例如，"这跟1971年世界锦标赛上鲍比·费舍尔（Bobby Fischer）的攻击路数有点儿像"。由于象棋有无数种可能的对弈状况，因此永远都存在不符合任何一种模型的状况。遇到这种情况，大师就会绞尽脑汁，努力为了一个具体的棋位而在先制和反制的决策树上寻找一种演绎性的解围方法。在演绎之树上，大多数顶级象棋棋手能够走到第三个层次。

人类的归纳模型认知和少量演绎逻辑混合的方法与IBM公司象棋棋手深蓝计算机的强攻型演绎方法形成了鲜明的对比。深蓝在一秒钟之内至多可以预估出两亿种可能的棋步，并设计出六层棋步和对抗棋步。深蓝的例子也说明了为何我们青睐归纳胜过演绎。演绎只能在定义十分明确的问题上（例如象棋）发挥作用，要想使用演绎法，问题不能存在任

何信息缺失或模糊。因此，演绎法是一种强大的推理方法，但具有天然的弱点。归纳法虽然更容易出错，但它也更具灵活性，能够更好地适应世界抛给我们的不完整和模糊的信息。从进化的角度来说，我们应该按照这个方向演变。深蓝的象棋水平或许能赶上加里·卡斯帕罗夫（Gary Kasparov），但加里·卡斯帕罗夫的归纳机制让他不仅能够下象棋，还能在早晨系好鞋带，在餐馆点餐，而这些任务则会让深蓝完全不知所措。[34]

现代认知科学关于人类行为的观点几乎与传统经济学的完全理性观点完全相反。[35]完全理性假设人类具有100%的演绎性能，跟深蓝一样，永远都在处理非模糊的、定义明确的问题。它还假设我们不会学习，毕竟，如果我们已经是完美的了，又有什么学习的必要呢。

认知科学的观点还为实验经济学家遇到的异常结果提供了一种解释。框架问题、可用性偏差、锚定以及其他效应都符合人类作为反应快速的模型识别者和模型完善者的形象。有时候，我们会在匆忙的归纳中犯错误，逻辑遗漏接连发生。进化在让我们变得快速、灵活、大多数时候正确，而不是缓慢、脆弱、在逻辑上保持完美。但这也让我们陷入了一种困境。完全理性的主要益处在于它很具体，我们可以围绕它写出一系列方程、构建模型。经济学要想成为科学，就需要这样的精准，那么我们用模式识别、归纳、学习以及类比等含糊不清的理念怎么能够完成这个目标呢？

行为主体的大脑

当前尚未出现标准的、广泛认可的归纳推理模式，但许多研究者已经证明，建立准确表述、数学化的归纳模式至少是有可能实现的。与模式识别以及学习相关的模式已经成为计算机科学研究的主要任务，许多这样的模式都已得到实际应用，从在机场识别恐怖分子的面孔到识别信用卡欺诈都包含在内。

来自密歇根大学的计算机科学家约翰·霍兰德，来自加州大学洛杉矶分校的心理学家基思·霍利约克（Keith Holyoak），来自密歇根大学的心理学家理查德·尼斯贝特（Richard Nisbett）以及来自普林斯顿大学认知科学实验室的保罗·萨加德（Paul Thagard）已经设计出了一个通用的认知模型，它展示了新经济人的可能面貌。[36]霍兰德及其同事所构建的模型的基本结构包含以下构成：

- 行为主体。有一位行为主体与其他行为主体及所处的环境进行互动。
- 目标。行为主体拥有一个或多个试图达成的目标，这样行为主体就能感知到他的当前状态与目标状态之间的差距，例如，“我很饿”或“我遇到危险了”。行为主体的工作就是作出更加接近目标的决策。
- 经验法则。行为主体拥有一些经验法则，能够根据世界的当前状态制定行动计划。这些法则被称为“条件-行动规则”，或者说更广为人知的“如果-那么规则”。举个例子，“如果火炉很热，那么不要触摸”。行为主体在任何时候的经验法则集合被称为行为主体的心智模型。
- 反馈与学习。行为主体的心智模型记录了哪些规则曾帮助他实现目标，而哪些规则曾让他远离目标。曾经取得过成功的规则的使用频率通常高于失败的规则，从环境中获得的反馈能够让行为主体不断学习。

归纳本质上是一个解决问题的工具，行为主体通过这个工具来接近目标。规则的集合会受到行为主体所处环境的反馈影响，为行为主体的外部世界创造出一个内在模型。随后，行为主体会利用这个内在模型来预测应该采取何种措施以便更好地应对在追求目标过程中所遇到的各种形势。

这种归纳式地解决问题、学习以及通过类比进行推理的过程看起来相当复杂。然而，正如霍兰德所言，这种机制在生物界普遍存在，只不过其复杂程度各有不同。低级的细菌也会使用归纳式的问题解决办法——在遇到不同浓度的食物时，细菌会朝着浓度更高的方向前进，这种行为中隐含了一个预测，即那个方向会有更多它所喜爱的微粒。细菌的DNA提供了一个细菌世界的心智模型：如果食物的浓度正在增加，那么很有可能会继续增加（如果食物分子是完全随机分布的，那么这个结论不正确）。细菌拥有目标（食物）、识别模型（化学梯度），会进行预测（食物在哪个方向），并且采取行动（摆动鞭毛）。随后，细菌会从环境中得到非常直接的反馈：如果规则好使，那么它就能存活和繁衍，否则就会死掉。

青蛙的学习

假设有一只青蛙名叫克米特，它幸福地坐在池塘中的睡莲叶子上。[37]它的蛙生目标非常简单：远离危险，吃到苍蝇。克米特的心智模型里拥有各式各样的探测器，能够通过感官来检测环境中的状况，比如移动的、有条纹的、大型的、近处的、嗡嗡作响的。克米特还有一套描述它所能进行的各种行为的感受器，例如逃跑、求偶、伸舌头、无动于衷。克米特的心智模型所要做的工作就是将探测器和感受器一起放进“如果-那么”规则里，以便更好地实现它的目标。举例来说，克米特的大脑中或许有这样一条规则：如果遇到小型的、飞行的、位于视野中心的物体，那么就要伸出舌头。

克米特在最开始的时候是如何获得这些规则的，它又是如何不断学习的呢？有一套基本标准是通过DNA嵌在克米特体内的，因此在蝌蚪时期，它就拥有能够进食和躲避危险的初始设置。然而，一旦进入广阔世界，克米特的规则设置就会马上开始扩展。比如，克米特身上与生俱来的规则是“如果遇到是小型的、飞行的、位于视野中心的物体，那么伸长舌头”。但在经过几次与蜜蜂和马蜂的不愉快接触后，它或许会将规则修改为“如果遇到小型的、飞行的、蓝色的、位于视野中心的物体，那么伸出舌头”，以便更准确地捕捉到苍蝇。同样，由于蜜蜂和马蜂带来的糟糕经验，它或许会增加一条规则，“如果遇到小型的、飞行的、有条纹的物体，那么不做动作”。

随着规则根据环境所提供的反馈不断修改和扩展，它不可避免地会遇到矛盾。我们可以想象有一天，克米特看见了一个小型的、飞行的蓝色物体，与此同时，一片阴影掠过头顶，那么它的探测器会激活两条规则“如果小型的、飞行的、蓝色的物体位于视野中心，那么伸出舌头”；另外，“如果阴影在头顶，那么逃跑”。这两条规则都在争夺克米特的注意力。我们将会假设，为了解决这个矛盾，克米特的心智模型进行了一个可信度赋值的过程，它会根据过去这两条规则在实现目标过程中的表现来给它们打分。成功的规则能够得到高分，失败的规则则会得到低分。“如果阴影在头顶，那么逃跑”这条规则对于帮助克米特避免成为猛禽的午餐发挥了重要作用，在克米特的心智模型的打分系统里，这条规则会被赋予100分。与此同时，捕捉苍蝇的规则也很重要，但是不如前面那条规则重要，因此它会获得90分。这样一来，两条规则就都被激活了，但克米特的心智模型将会选择得分更高的规则，然后逃跑。接下来，它的心智模型还会记下规则是如何为目标服务的，并且更新每条规则的得分。

这个过程自然而然地导向了学习。比如，克米特有一条这样的规则，“如果听见‘嗡嗡’声在左边，那么向左转头”。但偶然有一次，克米特把头转向右边并碰巧捉到了一只苍蝇，于是它又有了一条规则，“如果听见‘嗡嗡’声在左边，那么向右转头”。这两条规则是矛盾的，且都在过去让它获得了午餐，所以这两条规则各得10分打平。现在，我们可以将这两条规则比较一下，在青蛙的世界里究竟哪条规则可以带来午餐。只要分数相等，克米特就都会试一试。我们会清楚地看到，一段时间过后，向左转头规则的效果好过向右转头，它的分数也会不断变高。

然而，这其中存在两个问题。首先，我们不希望这个系统变得过于死板——一旦某条规则的分数从10变成了11，克米特就永远不会再尝试其他的规则。毕竟，如果克米特尝试向右转头，然后偶然间捕捉到了一只苍蝇，让该规则的分数变成了11，死板会导致它永远不会尝试向左转头规则，然后被饿死。因此，更高的分数可以增加规则被激活的可能性，但不确保一定会激活。我们还希望用一些分数非常接近的规则进行实验，甚至会希望偶尔尝试不受待见的规则，看看世界是否发生了变化。其次，我们还会遇到一个问题——行动和反馈有时候是分离的。这样的系统如何能够学会依照战略行动呢？也就是说，在短期内会有成本而在长期有回报的时候，克米特会怎么做？

霍兰德及其同事的答案温暖了经济学家的心——我们拥有市场。他们的假设是，存在于你心智模型里的规则会互相竞争，利用各自的信用分数竞标你的注意力。[38]信用分数越高，一条规则胜出的可能性就越大，但并不一定胜出。比如，竞标规则之间的5分差距或许意味着它们被选择的概率是50：50，而10分的差距意味着60：40的概率，以此类推。此外，竞标通常不是由规则本身，而是由复杂的规则链决定的。举个例子来说，克米特或许有这样一条策略——寻找苍蝇很多的地方并接近（暂时会吓走苍蝇），静坐一会儿直到又有苍蝇飞回来，然后用舌头击中它们。这个过程中所涉及的复杂规则链说明，获得回报的规则（如果出现苍蝇，那么伸出舌头）与规则最初被设立时的情况（如果有有气味的东西吸引苍蝇，那么靠近）存在一些差别。

在规则市场中，我们能将链条上的这些规则想象为供应商和消费者之间的相互作用。比如，“伸出舌头”规则购自“静坐不动”规则，“静坐不动”又购自“靠近有气味的物体”规则。因此，当伸出舌头获得回报时，它必须向供应商规则支付，供应商又会向上一层供应商支付，等等。这样一来，创造利润的规则（它们增强了规则链获得回报的能力）会不断增

强，被使用的频次也会增加。然而，随着规则不断向链条的远处回溯，支付的涓滴效应也会不断减小，这可以归因于中间商的存在。由此产生的更低支付为原因和结果之间的距离设定了自然而然的限制。这样的结构与实验假设相符，即在某种程度上我们能够依照策略行动，但在很长、很复杂的因果链中就很难进行推理。

青蛙的诗？

由此，我们可以看到由竞争性规则、分数和环境反馈所组成的相对简单的系统是如何形成灵活的、能够不断学习的模式识别系统的。我们将再次进行假设，看看系统是如何大幅度提高性能的。假设随着时间的推移，系统中的规则将通过自组织形成不同的阶层。由于克米特的世界存在规律性，在克米特的心智模型规则里，最先被激活的模型也具有规律性。比如，与小型的、飞行的蓝色物体相关的规则往往会同时被激活。如果某些规则经常同时被激活，那么它们就会相互关联，我们可以认为它们被分到了一个类别。克米特遇到小型的、飞行的物体时所激活的各种规则可以被归到苍蝇类别里；同样，激活大型的、飞行的、扑翼规则的物体可以被归到鸟这个类别里。克米特的心智模型里存储了许多经验，会相互关联着对苍蝇和鸟作出反应。

这种层级结构赋予了归纳性系统两个重要的优势。第一个优势是结构让系统能够对新事物作出反应。任何心智模型都不可避免地会比它所在的世界简单。因此，克米特永远都会遇到它从未经历过的状况，需要作出一些反应（哪怕是无动于衷）。因此，我们假设克米特大脑中的规则是按照默认层级设置的。当克米特遇到一种情况，它的心智模型就会进行扫描，看看有无针对这一情况的具体反应；如果没有，它就会求助于默认反应。[39]举个例子来说，克米特或许拥有如下默认反应：如果物体移动且没有其他信息，那么就要逃跑。这是一个通用、保守的规则，其目的是让克米特摆脱困境。然而，在这种通用的默认规则之下，克米特或许会遇到多种情境，拥有更多信息以及针对特定情境的反应，比如，如果附近的小型物体在移动，那么缓慢地靠近。因此，克米特在遇到新事物的时候会比较谨慎，会求助于靠得住的规则，但一段时间之后，它将针对新情况形成一套调整后的技能。

规则层级的第二个优势在于它可以通过类比进行推理。[40]青蛙或许无法想出字面意义上的隐喻（因此不存在青蛙诗歌），但使用归纳系统

的青蛙或其他行为主体可以通过类比进行推理，它们可以说某事物与其他事物相似。层级中的规则集合自然而然地将它们引向了这类推理。比如，克米特拥有的与鸟相关的大部分经验都是跟海鸥打交道，它拥有一个分类叫作鸟，在探测到大型的、飞行的、扑翼的东西时能被激活。有一天，克米特看见了某种大型的、飞行的但并非扑翼的物体，而这种相遇会激发它部分而非全部的探测器。对于克米特来说，这种新物体“跟鸟类似”，但并不完全符合与鸟相关的经验。克米特的心智模型会扫描其他类别以寻求匹配，结果发现它与鸟类的相似性胜过苍蝇与狗。因此，克米特会利用自己的默认层级，搜寻最通用的相关反应，也就是逃跑。通过类比进行推理的能力让处于模糊世界的、不断遇到新事物的克米特拥有了一个很大的优势。相对于将新的不明物体归类于“我不知道它为何物”，当看见某物体跟鸟类似时，克米特更有可能产生正确的反应。举个例子来说，如果克米特看见的非扑翼物体是一只翱翔的、正在寻找青蛙当点心的鹰又会怎样？这个物体还有可能是飞机或其他无害物体，但将它看作与鸟类相似的物体依然是合理的，可以激活保守的反应。关键在于，即便反应并不完美，它作出合适举动的概率也会提高。随着时间的推移，如果鹰或飞机的概念进入克米特的世界，那么它的“跟鸟类似”的定义就会进化，从而在默认层级里发展出更加合适的反应。

股票机器人

我们可以看到个人如何创建青蛙捕捉苍蝇的模型，那霍兰德及其同伴的归纳模型应该如何应用到人们的经济决策过程中呢？在1987年圣塔菲研究所的经济会议后不久，霍兰德开始与布莱恩·阿瑟合作。尽管阿瑟是一位学者，但他在职业生涯的早期就投入了商业领域，攻读硕士学位期间在麦肯锡公司待过一段时间。他亲身见识过现实世界里的决策过程，它与经济模型里冷冰冰、干巴巴的完全理性截然不同，因此他对于寻找一种更好的表达方式产生了持久的兴趣。霍兰德与阿瑟决定组成团队，将霍兰德的理念应用到股票市场行为的模型构建中去。他们同布莱克·勒巴伦（Blake LeBaron）、理查德·帕尔默（Richard Palmer）、保罗·泰勒（Paul Tayler）共同构建的“圣塔菲研究所人工股票市场”模型已经成为复杂经济学文献里的一个经典。[41]

跟通常的归纳模型一样，圣塔菲研究所模型里的行为主体被视作信息处理的实体，它们本质上是小型的电脑程序。这些行为主体会从周围的环境中收集信息、处理信息、产生决策，然后接受环境对这些决策的

反馈。

与糖域中的行为主体一样，人工股票市场模型里的行为主体也会受到进化力量的影响。就像大多数取得经济成就的糖域行为主体能够生存和繁衍一样，在圣塔菲研究所的股票市场模型中，最成功的股票交易行为主体会得到回报，而最落魄的行为主体则会破产和出局。阿瑟、霍兰德及其同事在他们的模型中增加了第二层次的进化。在人工股票市场模型中，行为主体的大脑也在发生进化，使得它们能够学习。

为了弄清其中的过程，我们可以假设你买彩票中了一笔不大不小的奖金，然后希望将它用在投资上。你给股票经纪人打电话，寻求一些关于投资策略的一般性建议。经纪人为你提供了一些经验法则，例如“买入并持有是最佳的长期投资策略”，“永远要确保你的投资组合是多样化的”以及“注意市盈率远高于同领域其他股票的股票”。不过，你决定多找几位经纪人问问。你又给其他几位经纪人打了电话，还有你的叔叔以及几位你觉得应该问一问的朋友。很快，你得到了一堆关于投资的建议，其中有一些是相互矛盾的。你的下一步工作是要将浮在脑海中的这些潜在策略整理一下，看看下一步该如何行动。

在这个场景中，市场上存在两个级别的竞争。首先是投资策略的一般性竞争：成长型投资者和价值型投资者、牛市和熊市，究竟谁的表现更佳？其次是你脑海中相互矛盾的潜在投资策略对于你注意力的争夺：在经纪人和叔叔赫比之间，你应该听从谁的建议？阿瑟和霍兰德推测，这两种竞争将会遵循进化式的学习过程。

为了探究这一观点，阿瑟、霍兰德及其同事在电脑上搭建了一个模拟交易环境，里面只有一支股票，而且分红是随机的。随后，他们又创造了100位进行交易的行为主体，行为主体可以买进和抛售股票。股票的价格由市场上的行为主体的买卖决定。每一位行为主体都有一个简单的目标，那就是尽可能多地赚钱，为了实现这个目标，每一位行为主体都要决定何时买入以及何时抛售。行为主体们可以获得三类信息以便进行决策：股票的历史价格模型、股票的历史分红以及无风险利率。接下来，阿瑟、霍兰德及其同事就要决定如何让行为主体处理信息并作出决策。

跟青蛙的案例一样，研究者们使用条件-行为规则将市场中的模型映射到行为主体对股票价值的预期上。比如，有的规则可能是“如果价格在上一周期上升5%，那么预计下一周期的价格将在本周期的基础上再涨5%”。遵守这一规则的行为主体仅仅通过比较预测价格与当前价格就能

决定是买还是卖；如果预测价格上涨，行为主体将买入；如果预测价格下跌，行为主体将抛售。这样结构的规则也有可能被设置得非常复杂："如果价格在过去三个周期中都在上涨，并且价格不高于股息与无风险利率之商的16倍，那么预测下一周期的价格加上股息等于当前价格加上股息的106%。"行为主体所使用的经验法则可以基于基本原理（例如在某个市盈率范围内买入股票），或基于市场趋势（例如在价格上涨时买入股票），抑或是基于两者的综合。在模型中，规则的复杂程度不受限制。

就像我们假设中彩票的人要对相互矛盾的股票建议进行整理，模型中的每一位行为主体都不只拥有一条经验法则——阿瑟和霍兰德为每一位行为主体配备了100条经验法则。在对比哪条假设能够带来股市上的成功时，行为主体的软件"脑袋"里会出现100条规则。也就是说，每一条规则都是一种潜在的投资策略，而行为主体的工作是要厘清属于他的一套潜在策略，然后找到哪些策略能够帮助他赚钱。接下来的问题是，行为主体是如何对他的100条相互抵触的投资策略进行分类的呢？答案相当简单：跟青蛙克米特一样，行为主体利用的是在过去起作用的那些规则。

阿瑟和霍兰德设置了一套在行为主体大脑内发挥作用的进化流程。每一条条件-行为规则都被编码为一连串的1和0，我们可以把这些字符串当作电脑的DNA，它们代表着各不相同的投资策略。因此，011000100101有可能代表着"如果在过去5个周期中，股息平均下降了10%，那么预测价格将在下一周期下降2%"，而100011010101有可能代表着"如果无风险利息高于股息3倍，那么预测价格将在下一周期上涨5%"。如此一来，每一位行为主体在其软件大脑里都有一堆投资策略DNA。每一个策略DNA都被赋予了一个适合度评分。如果某个策略让行为主体赚到了钱，它的适合度评分就会增长；如果某个策略导致行为主体亏了钱，它的适合度评分就会降低。

在每个回合的游戏中，各个行为主体都会经历以下流程。行为主体被告知关于历史股价、股息和无风险利息的信息，随后，行为主体将这些信息与潜在策略和模型仓库进行对比，寻求匹配。如果可匹配的规则不止一条，行为主体将会查看分数并选择分数更高的那条。因此，在过去行之有效的规则的利用率往往更高。在应用规则后，行为主体将会观察它是如何发挥作用的——是赚钱还是赔钱？如果行为主体能够赚到钱，就会进一步强化规则；反之，规则将被削弱。

至此，我们就拥有了一个可以通过固定数量的规则进行学习的系统。假设每位行为主体起初都拥有100条随机构建的规则。行为主体使用上述系统，经过试验和犯错，最终会学习到哪些规则能帮他赚钱，哪些不能。但在现实的股票市场中，人们一直在不断创造新的投资策略，那么我们该如何让模型行为主体进行创新并创造新的规则呢？

阿瑟、霍兰德及其同事在模型中构建了以下流程。在并不一致的时间周期中，每位行为主体的100条规则将会经历一个进化流程：表现较差的20条规则会被消除，新的规则会被创造出来以填补空位。表现更好的另外80条规则的电脑DNA的单个元素会发生基因突变。比如，某个1会随机变成0，或者反过来。其余的某些策略则会以“电脑婚配”中进行重新组合。某个DNA字符串会在某个随机的时间点分裂成两份，并与其他分裂成两份的字符串结合，从而产生新的字符串。基因突变和重新组合的作用在于为行为主体的100条规则池创造出新的策略。许多新规则是无厘头的，甚至是有害的，有些则是成功的创新，甚至比其他在市场上发挥作用的策略还要成功。

随着归纳式股市准备就绪，阿瑟、霍兰德及其同事启动开关，进行了一系列实验。在运行完全模型（每位行为主体拥有100条规则）之前，他们进行了一次实验——每位行为主体都只有唯一一条相同的规则（即完全理性），并且学习能力为0。[42]这轮实验的结果与传统经济学的预测十分相似。模型很快就在一个接近于理论均衡价格的数值上稳定下来，这个价格与股票的基础价值相当。此时只有很少的交易量或波动，没有人能比其他人赚到更多的钱。这个实验表明，如果每个人都按照（或接近于）传统理性模型行事，结果就会与传统理论所预测的相差无几。

随后，他们又进行了第二轮实验——激活每位行为主体“脑中”的100条规则，在初始化时为这些规则赋予少量随机的策略，并且将学习能力提高到0以上。这些设置对模型的行为产生了巨大影响。随着时间的推进，交易量上升，波动幅度增大，股价的动态变得更加复杂，其中还包括泡沫和崩盘。市场状况也变成了一段时间静止与一段时间风雨满楼相交织的模式。此外，行为主体的相对表现也出现了巨大差异，部分沃伦·巴菲特式的、表现超级优异的行为主体开始展露光芒，而另一些则陷入了破产的泥潭。正如我们将在第17章看到的那样，真实的金融市场与这种景象更为接近，而不是像传统经济学所预测的那样静止。

结果变得更加动态、更加接近现实市场的原因如下。如果每一位行

为主体在开始时拥有同样的完全理性，并且由于不会学习一直保持着这种状况，那么市场就会缓慢向前，价格大致维持在均衡状态。一旦引入了规则的多样性以及学习，情况就会变得更加复杂。假设因为某些原因，股票价格开始上涨。行为主体中的部分成员拥有积极的规则，他们寻求股票价格的上涨然后买入。当价格上涨，就会有越来越多这样的参与者涌入，造成价格越来越高。而另外一些参与者会关注基本价值，到某个点就会抛售，因为他们认为股票估值已经过高。如果有足够多这类行为主体瞬间涌入，就会导致股票价格低迷，从而激发那些遵循在股票价格下跌时抛售规则的参与者涌入。随着越来越多的参与者想逃离出场，股价就会快速下跌。新的规则甚至有可能进化到寻求这种起起伏伏的模型，以便从中获利。价格变动是由不同规则的动态互动而驱动的，与股票潜在经济价值的关系甚小，甚至毫无关系。复杂模型也不是仅仅由于随机噪音而导致的。行为主体的脑中时刻进行着一场关于信念的复杂斗争，这才是导致市场波动和复杂模型形成的根源。在第17章，我们将进一步探索不同类型的行为主体之间如何竞争，研究他们的策略或许会对市场行为的某些重要方面产生影响。

在经济的复杂适应系统中，理解个体的微观行为对于理解系统的整体行为非常关键。在过去100多年的时间里，经济学家都在用他们的人类行为模型勉强对付——这在今天的大多数经济学家看来都是过于简单、在根本上存在矛盾的，仅仅是为了在数学上容易处理。如今，物理学家、计算机科学家和认知科学家的合作产生了一种新的经济人模型。这种新模型将人类描绘成归纳式的理性模式识别者，他们能在模糊和快速变化的环境中作出决策，并且能够不断学习。同样，真实的人既不是纯粹的利己主义者，也不是纯粹的利他主义者。他们的行为是为了激发社交网络中的合作行为，奖励合作者并惩罚“搭便车”的人。我们当中没有人是完美的，任何人都有着缺点和偏差。

然而，还有人对完全理性进行着最后的辩护。有人或许会说，只有一条道路可以通往完全理性，至少经济学家有着一个标准的、广受认可的方法来为行为构建模型，哪怕它是有缺陷的、简化的。归纳式理性甚至还做不到这一点。认知科学仍然处于婴儿期，尚未找到一个广受认可的方法。尽管如此，我仍然认为，使用一个与我们所知的事实相符的模型胜过于使用已知是错误的模型。此外，世界上也许永远不会存在一个

能够满足所有目的的经济人。股票市场所需的简化版行为假设也许同微观经济学模型所需的不一样。在这里，我想重申的一点是，如果有人认为理论好比地图，那么抵达经济人的不同道路至少应该是互不矛盾的，并且与我们所知的真实人类行为一致。与此同时，认知科学将会继续按照其自身的节奏快速发展，建模技术将继续提升，而经济学永远不会成为“心理历史学”，我们将对行为如何驱动经济这一问题获得比如今更深刻的理解。

07

视角三：经济是一个复杂互动网络

对于任何复杂适应系统而言，网络都是一个关键因素。若行为主体之间没有互动，就不存在复杂。比如，生物世界是由一个庞大的网络层级组成的：分子在细胞之中互动，细胞在生物体之中互动，而生物体在生态系统之中互动。人体是由互动系统组成的高度复杂的网络集合体，这些系统包括大脑、神经系统、循环系统和免疫系统。如果将人体的网络系统去除，那么每个人都不过是一小箱化学物质再加半浴缸水。

经济世界同样依赖于网络。地球被道路、下水道、供水系统、电网、火车轨道、输气管道、无线电波、电视信号以及光缆环绕着。这些东西提供了物质、能源和信息的高速通道和偏僻小径，它们在经济的开放系统中川流不息。此外，经济还包含着大规模的、复杂的虚拟网络：人们在公司之中互动，公司在市场之中互动，而市场在全球经济之中互动。跟生物界一样，经济世界的网络是由网络之中包含网络的层级构成的。

尽管网络对于经济活动来说极其重要，但它们一直没有成为经济学家关注的重点，而社会学家在近年来对网络做了许多研究，但通常来说，他们是在社会政治关系的背景下而非经济学的背景下进行研究的。[1]传统经济学不愿意把网络考虑进来，这是因为网络无法很好地融入均衡的范式。传统经济学模型通常会假设行为主体只在拍卖活动（或者其他一些定价机制）或一对一的议价中进行互动。这种假设之所以产生，是因为拍卖和一对一议价可以被描述成均衡系统，而一大群人进行复杂互动很难用数学进行建模，往往需要借助计算机模拟。

多年来，网络已经成为物理科学家最感兴趣的话题之一。20世纪50年代到60年代，匈牙利数学家保罗·厄尔多斯（Paul Erdôs）与阿尔弗雷德·瑞利（Alfred Rényi）做了许多开创性工作。[2]近年来，新的数学工具和计算机极大地加速了物理科学和社会科学对于网络的研究。研究表明，网络具有许多通用的特性，这些特性无论是在粒子互动的网络、大脑的神经网络，还是组织之间的人际网络中都是适用的。[3]在本章，我们

将会看到最新发现的一些“网络法则”在经济学系统中意味着什么。

在探索网络的特性之前，我们应该先对一些术语进行定义。拿一张纸，在上面画几个点，然后将这些点用线连接起来——这样你就能画出一个网络。数学家将这些点称为节点，称连接它们的线条为边。网络本身的概貌被称为图形。如果在连接这些点的时候，你是随机地画出线条，那么你所画出的网络就被称为随机图（见图7-1）。如果你所画出的连接线条是规律性的模型，比如每一个点与它相邻的最近的四个点相连，从而形成棋盘状图案，这样的图案被称为栅格图。随机图和栅格图在经济领域都有体现，某些最有趣的网络就是两者的综合体。

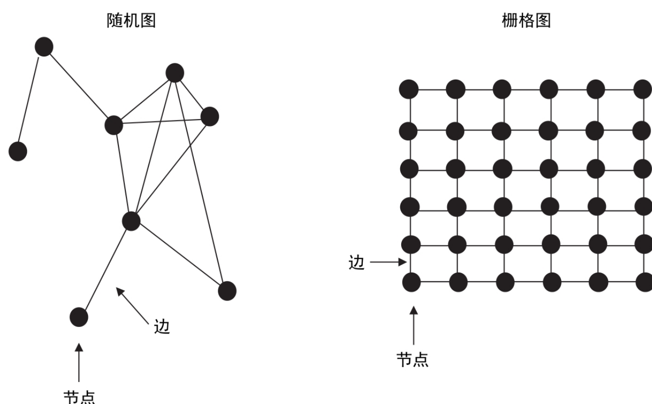


图7-1　随机图和栅格图

网络是如何变得火爆的

经济学家已经意识到，某些产品，例如电子邮件、传真和电话，拥有一个共同的特征，那就是使用这些产品的人越多，它们的用处也就越大，这被称为网络效应。然而，传统经济学在历史上并没有过多地阐述为何这类产品会在人群中突然火爆并流行开来。

数学生物学家斯图尔特·考夫曼认为厄尔多斯与瑞利所开发的随机图理论中蕴含着上述问题的答案。[4]请想象1 000枚纽扣散落在硬木地板上。假设你的手里有一些线，你随机拾起两枚纽扣，用线把它们连起来，然后又把它们放回去。在最开始的时候，你拾起的每一枚纽扣大概率都是没有跟别的纽扣连在一起的，就这样，你能让许多纽扣两两相连。随着工作的不断进行，在某个时刻你就会拾到一枚已经与其他纽扣相连的纽扣，然后你会再加上第三枚。最终，没有被连接的纽扣越来越

少，一些有四枚纽扣或五枚纽扣的纽扣串开始形成，就像一座座小岛散落在纽扣的海洋里。

接着，随着你不断串联纽扣，孤零零的纽扣串突然开始连成巨大的超级串——两个5枚的串连成一个10枚的串，一个10枚的串加上一个4枚的串成为14枚的串，以此类推。物理学家将系统的某个特征突变称为相变。举例来说，每次将蒸汽的温度降低一度，到达100℃时，蒸汽就会突然变成液体，而到达0℃时会变成冰。在随机网络中，从小型串到巨型串的相变发生在某一具体时刻，即线段（边）与纽扣（节点）的比率值超过1，也就是说平均一条线段对应一枚纽扣。[5]我们可以认为一条线段对应一个节点的比率是一个“引爆点”，此时，随机的网络突然从稀疏连接变成了密集连接。[6]

考夫曼认为，网络形成过程中的这个“引爆点”有助于解释生命所需的化学反应网络最初是如何形成的。我们也在经济和科技环境中发现了这种效应，互联网或许就是一个最新案例。互联网是在20世纪60年代由美国发起的一个国防项目中被发明的。它坐了20年的“冷板凳”，当时主要用于学术领域。后来，到了20世纪90年代，互联网的应用突然之间爆发了。有一种合理的解释是这样的：更快、更便宜的调制解调器以及更优的用户界面将人们的社交网络的节点比率值提升至了神奇的1，从而带来了互联网用户的爆发式增长。

随机图理论是一种能让我们用模型表达此类现象的方法，从中还可以看到网络效应并不是平滑和渐进的，而是高度非线性的。通过分析此类网络，我们或许能够更好地理解为何有些时尚能够风靡，某些政治运动是如何在突然之间爆发的，以及是什么导致了股市的波动。[7]

世界很小

1967年，心理学家斯坦利·米尔格拉姆（Stanley Milgram）决定开展一项实验。[8]他向居住在堪萨斯州和内布拉斯加州的一群人写信，并说明须将这些信转发给居住在波士顿的两位收信人中的其中一位，但他只指明了收信人的姓名、职业、人口统计资料和大致位置。米尔格拉姆指示堪萨斯州和内布拉斯加州的参与者将他们收到的信转发给一位熟人，这人又将信转给自己的一位熟人，直到信到达计划中的收信人手里。如果你是参与者之一，你或许在波士顿不认识什么人，但你将信转给了在美国东海岸上大学的俄亥俄州的表亲，他又将信转给了一位住在波士顿

的大学同学，同学又将信转给了一位医生——因为计划中的收件人是一名医生，等等。米尔格拉姆惊讶地发现，大部分信件最终都顺利抵达，并且发信链上涉及人数的中位数是6。这个结果成为“六度分割理论”的基础，该理论的主要内容是指地球上的每个人都可以通过6个人（或更少）联系到任何人。

来自哥伦比亚大学的邓肯·瓦茨（Duncan Watts）[(10)]与来自康奈尔大学的斯蒂芬·斯托加茨（Steven Strogatz）[(11)]通过“凯文·贝肯游戏”，提出了更新的、异想天开的“六度原则”。[9]据瓦茨介绍，这个游戏是由威廉与玛丽学院的一群兄弟会成员发明的，他们都是资深影迷，于是决定让凯文·贝肯成为游戏世界的主角。[(12)]在游戏中，每个人都可以想象一个随机的演员，然后决定他/她与贝肯电影之间联系最少需要几个步骤。比如，凯文·科斯特纳（Kevin Costner）的贝肯数为1，因为他跟贝肯一起出演了《刺杀肯尼迪》；布鲁斯·威利斯（Bruce Willis）从未与贝肯一起演过电影，但他的贝肯数是2，因为他与帕特里克·麦科马克（J. Patrick McCormack）一起出演了《世界末日》，而麦科马克与贝肯一起出演了《透明人》。通过使用弗吉尼亚大学布雷特·查登（Brett Tjaden）的数据，瓦茨及其同事发现，美国演员的最高贝肯数是4。随后，他们又对互联网电影数据进行了彻底的研究，发现世界上大约57万演员中90%的人与贝肯或多或少都有关系，世界上最大的贝肯数是10，其中85%的人的贝肯数小于等于3。[10]瓦茨与斯托加茨在其他社交网络中也进行了类似的研究，包括科学家和企业董事会成员，也得到了类似的结果。归根结底，我们只能说这个世界很小。[11]

地球上有70多亿人口，我们怎么可能与其他人之间只隔着6个人呢？数学家或许会指出，如果世界上的每个人平均拥有100位朋友，那么每个人的朋友的朋友的数量就是10 000（100×100）个，朋友的朋友的朋友数量为100万个，以此类推，在第五层，我们就可以连接到10亿人。[12]但这其中存在两个问题。首先，地球上每人平均的朋友数量为100位，这个估值恐怕是过高的。其次，人们的朋友圈会有相当程度上的重合。也就是说，我的100位朋友与我朋友的朋友并不是完全不同的，因此，所谓的小世界效应必然是由其他原因造成的。

瓦茨和他来自密歇根大学的同事马克·纽曼（Mark Newman，他们都是圣塔菲研究所的成员）给出了一个有趣的答案：小世界效应是由网络本身的结构造成的。瓦茨和纽曼发现，社交网络已经进化成为规律性与随机性的高效混合体。[13]假设有一张美国地图，我们用小圆点来标记

所有人口数量大于10万的城市，用线条把一座城市与距离它最近的四座城市连接起来。比如，波士顿可以跟伍斯特、坎布里奇、普罗维登斯和曼彻斯特连接起来。这样一来，我们就创造了一个栅格网。这样的栅格网或许看起来不如图7-1那样整齐，但“四座距离最近的城市”规则意味着该网络具有规律性和结构。我们可以观察与波士顿相连的城市，并将它们看作一个区域，比如，新英格兰东南区。这种规律性的缺点在于穿越整个网络需要许多跳点。如果我们想把波士顿与圣地亚哥连接起来，就必须将波士顿连到普罗维登斯，再把普罗维登斯连到哈特福德、哈特福德连到纽约市，等等，每次连接都需要穿过整个国家。由于栅格图的分散性很高，所以从一个海岸到另一个海岸或许需要连接多达20或30个点。

现在，将最近城市原则放在一边，我们说每座城市与其他四座城市是随机地连接在一起的，从波士顿开始的跳点或许会经过纽约州的奥尔巴尼、堪萨斯州的托皮卡、佛罗里达州的萨拉索塔以及华盛顿州的斯波坎。有些随机线条肯定很短，例如华盛顿特区与费城之间的线条；有些是中等长度，例如从丹佛到克利夫兰；而其他的则相当长，例如从圣地亚哥到夏洛特。由于连接是随机的，短线、中等线和长线的数量将会相同，这就意味着连接任意两座城市之间的站点会少得多，因为对于任何两座城市而言，我们都可以根据跨度和距离选择较短、中等和较长的站点连接旅途。从栅格图换到随机图，分割度的数量变得屈指可数。

随机朋友的价值

社交网络就像栅格网，因为它们包含着秩序和结构。你的朋友中极大概率包含了你的发小、校友、同事、同行和现在的邻居。这就意味着，由于我们倾向于从社交池里找朋友，所以你的朋友们相互认识的可能性高于平均水平。所有社交网络都有着明显的同类人群或派系，比如圣路易斯的牙医往往都相互认识，斯图加特的飞机爱好者也一样。这些同类人群的存在证明网络并不是随机的，而是包含着秩序与结构。

尽管人们的社交网络是有结构的，但其中也会有一些随机的朋友——那些不在我们正常社交圈之内或在社交圈之外的人，那些我们在偶然情况下遇到并成为朋友的人，他们或许是你在度假时或在医生候诊室认识的。这些不在正常交往人群中的人是我们社交网络的桥梁，能够将我们连接到其他社交网络。如果你有一个结构良好的栅格图，然后再在

其周围放置几个随机的联系，你就能在拥有明显的同类人群的同时拥有很短的分割度，比如，“我的朋友大部分都是圣路易斯的牙医，但我在健身房认识的这位朋友在好莱坞工作”。这样的话，你很快就能跟麦当娜对话了。随机认识的朋友就像从夏洛特飞往圣地亚哥的特快航班。

纽曼和瓦茨的研究量化了这种效应。假设有1 000人，他们每人各有10位朋友以及0位“随机”朋友，也就是说每个人的朋友都是在严格定义的社交圈里结交的，那么我们就说平均的分割度值为50，即一个随机挑选的人连接到另外一人平均需要50个跳点。但如果现在改成每人的朋友中有25%是随机认识的，也就是说，有25%的朋友是假定的人在正常的社交圈之外结交的，那么平均分割度值就会降为3.6。[14]随机朋友这一理念与我们所认为的优秀网络建造者的概念在直觉上有些冲突，我们认定一个人拥有优越社会关系的前提往往是他对某一圈子有深入的了解。然而，瓦茨与纽曼的研究表明，社会关系最好的人实际上是那些与多样化群体有接触的人。我们都认识一些似乎能与任何人交谈，或在各种情境下与各行各业的人交朋友的人——他们才是真正的连接力最强的人。

社交网络的结构不仅对于个人而言十分重要，它还对大型机构的运行有着重大影响。如果机构为人们设定了严格的职业梯度，并且设有筒仓式的业务单元和部门，那么社交网络就会变得过于结构化而缺少足够的随机性，这也意味着信息需要经过长跳点链条进行传送，结果是沟通不畅、决策速度较慢。相反，有些机构会刻意地让人们不断在各种职能和业务部门之间转换，以此来创建公司内部的社交网络，以便拥有更加多样性的连接。人员的频繁变动可能会将社交网络变成随机噪声，但适度的人员变动则可以极大地改善社交网络的功能。比如，通用电气公司会让员工穿越不同的组织边界，并通过对不同地区的职员开展培训来锻造持久的社交联系。

“网络就是计算机”

多年来，太阳微系统公司（Sun Microsystems）一直在用“网络就是计算机”的口号为自己打广告。[15]这是一则非同寻常的、充满洞见的标语，因为它直指关于网络和计算机的基本真理。事实上，计算机就是网络，网络就是计算机。如果你拆解计算机，就会看到一堆用线连在一起的芯片形成了一个网络。你拆开其中一个芯片，就会看到数千万个晶体管通过线连在一起形成了网络。这些晶体管唯一要做的事情就是在两种

状态之间跳转：0和1。计算机不是从单个晶体管中获得动力的，而是从它们所组成的网络中获取。同样，单台电脑之间也可以串联成网络，从而创造出更强大的电脑。现代的超级计算机其实就是由单台电脑所组成的网络。

可以处于0或1状态的节点网络被称为布尔网络（Boolean networks），它是根据其发明者数学家乔治·布尔（George Boole）的名字命名的。[16]在布尔网络中，节点的0或1状态是由一套规则决定的。对此，斯图尔特·考夫曼用圣诞树上或开或关的闪烁串灯进行了解释。[17]假设只有三个灯串成一圈，它们分别是A、B、C。每一个灯泡可以是打开的也可以是关闭的，我们可以用0或1来分别指代这两种状态。每一个灯泡都可以收到两侧的灯泡信号，知道它们是开是关。假设每个灯泡基于从其他两个灯泡处得到的信号并遵循同一条规则来决定下一阶段的动作，灯泡A的规则是，如果灯泡B和C都是1，那么它也为1，否则为0（这就是所谓的布尔“与”规则，它是计算的基本构建模块之一）。如果A和B都是0，那么C也为0，否则C为1（这是布尔“或”规则）。我们为灯泡设定的规则决定了灯光随着时间的推移而闪烁的模式。网络在两种状态之间跳转，每个灯泡根据两个输入信号及规则，判断在下一个时间段是开还是关。由此，规则映射出两个时间段之间的网络状态。

从计算机芯片上的晶体管到化学反应网络（研究者甚至已经可以创建“化学计算机”），都存在着布尔网络。虽然我们的大脑使用的不是0和1的电脑逻辑，但个人的神经元可以通过类似的数学方式进行描述，因此，大多数研究者相信，大量的神经元就是某种形式的布尔网络。如果我们将经济视作一个庞大的大脑网络，那么实际上经济也是一个布尔网络（其复杂程度更加惊人）。无可否认，这有一点概念跳跃的意味，但如果是真的，我们可以推测这对于经济来说可能意味着什么。[18]

经过对布尔网络30多年的研究，人们已经充分地了解了它们的特性。尽管利用布尔网络可以做一些了不起的事情，比如组成万维网，组成你的身体和大脑，但它们实际上还是简单的创造物。此类网络的行为基本上是由三个变量引导的，第一个是网络中节点的数量；第二个是事物之间相互连接的程度；第三个是引导节点行为的规则中的“偏差”。

大即是美：信息规模

关于布尔网络的第一个重要事实是：网络所能拥有的状态数量是节

点数量的指数倍。拥有两个节点的网络可以有4种或者说2^2种状态：00、10、01和11。同样，拥有3个节点的网络可以有8种或者说2^3种状态。这一简单的事实导向了惊人的结果。如果一个网络拥有100个节点，那么我们用世界上速度最快的超级计算机将该网络的每一种状态都浏览一遍，需要5.68亿年才能完成。[19]如果再加5个节点（105个），那就超过了宇宙的寿命。如果我们无法探索这些小型网络，那么探索英特尔奔腾处理器或人类大脑的所有可能状态就毫无希望。[20]在所有这些可能的状态中，只有极小部分被探索过。这个问题的积极一面在于，随着网络不断扩大，它所能捕捉的信息量或者所能做的事情也会呈指数级增长。摩尔定律不仅使原始处理能力呈指数级增长，也使能力所能处理的事情呈指数级增长。

生物学能够很好地展示网络增长所具备的能力。基因组可以被认为是大规模复杂化学网络，它可以实现基因的开关状态转换。在启动“人类基因组计划”之前，科学家曾估计人类基因组大约包含10万个基因。但在基因网谱绘制完成后，他们惊讶地发现人类只有3万个基因。[21]相对而言，不起眼的蛔虫尚且有1.9万个基因，相当于人类基因数量的2/3。人类的复杂程度比蛔虫远远高出33%以上，但无比复杂的智人怎么会比简单的线虫只多出33%的基因呢？答案或许是，人类的基因是利用布尔网络来管理身体生长的，因此，虽然只多了1万多个基因，但人类基因网络所能产生的结果比蛔虫网络要复杂得多。

在任何一种信息处理实体的网络中，潜在的指数级增长创造了一种十分有力的规模经济。传统经济学常常认为，规模经济是一种与成本和数量相关的函数：如果部件的产量增加，那么生产每个部件所需的成本就会下降。布尔网络定律促使我们想到另外一种规模经济。随着布尔网络的不断扩展，创新的可能性也会呈指数级增长。由10个节点组成的布尔网络拥有2^{10}种可能的状态，而拥有100个节点的布尔网络则可以有2^{100}种状态。100个节点网络的可能状态不只是10个节点网络的10倍，而是比后者大了30个数量级（10^{30}）。从在当地的咖啡馆工作的10个人，到在波音公司工作的18万人，发生变化的不仅是员工人数扩大了4个数量级，而且复杂程度从制作拿铁变成了制造大型喷气式客机，其中的复杂程度提升了许多个数量级，就像人类的复杂程度比蛔虫高了许多个数量级一样。像波音公司这样大型的组织也内含着许多创新的空间——波音公司组织网络的大量状态意味着其赖以生存的可能方式要比街角咖啡店多得多。

如果说传统规模经济就是经济增长的全部，那么相对于200万年前，现在人类只不过是可以以更低的价格生产石器。但如果将人类组织视作一种布尔网络（我们承认其状态远不止开或关），我们就能看到，随着组织的规模不断扩大，创新的潜在空间就会呈指数级扩展。事实上，人类经济组织一直以来都在不断扩大。特别是，组织规模的跳跃式发展与技术的变革相吻合。定居农业的发展导致了村庄的出现，而村庄的规模远大于它的前一种组织形式——狩猎采集者部落。同样，工业革命催生了大规模工厂和工业城市，而20世纪晚期的信息革命催生了大量全球化企业。这些变化中存在一个良性的循环：技术变革催生更大规模的经济合作单位，这又将撬动更大的经济规模，最终创造出更多的未来创新的可能性。我们将在第三部分对此进行进一步研究。

然而，布尔网络却让我们在数学问题上陷入了困惑。如果说大型组织的创新空间大于小型组织，那么为何在商业历史上，小型组织的创新力往往比大型组织强呢？为何硅谷的小公司总是能打败那些巨人集团？

大即是糟：复杂灾难

网络理论向我们展示了关于规模的不同寻常且更加黑暗的一面。关于规模，还有一种重要的不经济性，它是由布尔网络的第二控制变量所驱动的：连接指数。你可以假设有一个网络，其中的连接非常稀疏，比如每个节点只与自己两边的单个节点相连，就像之前所说的圣诞树灯串。或者，你也可以假设一个连接非常密集的网络，其中有1 000个节点，每个节点都与其他所有节点相连，每个节点的连接数对于网络的特性有着重要影响。考夫曼及其在圣塔菲研究所的同事对这种关系进行了深入研究。[22]他们通过简单的观察得出了一个关键的结论：如果网络中每个节点的平均连接数大于1，那么随着节点数量的增加，连接的数量将会随着节点数量的增加而呈线性增长。这就意味着网络中相依性的增长速度大于网络本身，而这正是问题产生的根源所在。随着相依性的增长，网络某个部分的变化更有可能对其他部分产生涟漪效应。由于涟漪效应的可能性不断增加，网络某个部分的积极变化将导致其他部分出现消极变化的可能性增加。

为了说明这一点，我们可以假设你是一个小型创业公司的联合创始人，公司只有两个部门：产品开发部和市场部。你负责产品开发部，在有了一个关于新产品的点子后，你召开会议来讨论相关计划，市场部也

没有意见，你准备开始大展身手——就是这么简单。你的新产品大获成功，公司初具规模，你决定是时候设立财务部和客户服务部了。然而，跟所有的创业公司一样，你们公司有一点混乱无章，新建部门之间都互不沟通，员工都选择来找你。现在，你又有了一个新产品的设想，于是你跟市场部开会、跟财务部开会、跟客户服务部开会，以确保他们都会支持这项计划。这比之前要复杂一些，但还不算太糟糕。开会的总次数跟部门的数量一样，也就是说，开发第一个产品需要开一次会，开发第二个产品需要开三次会。

但你厌倦了当沟通枢纽，于是你告诉所有的部门负责人，他们应该直接沟通、分享信息、进行协调。很快，电子邮件开始漫天飞舞，会议室的日程应接不暇——你的促进交流的提议成功了。现在，你又有了一个关于第三代产品的设想，但是某些奇怪的事情发生了。你跟市场部进行日常会议，但在获得他们的同意之前，市场部经理告诉你，他们必须核对计划对预算的影响，而这需要财务部的批准。财务部那帮人说，在获得客户服务部提供的额外支持成本预估数据之前，他们不能批准你的项目。客户服务部需要跟市场部确认，以确保计划符合公司品牌和定价策略。突然之间，虽然公司的规模并没有发生变化，你需要参加的会议数量却从3场变成了10场（倘若所有的排列组合都会发生）。导致这个结果的唯一原因在于沟通连接点的密度增加了。如果公司继续扩张并且保持同样的流程——每个人跟每个人沟通，那将会是什么情况？如果再增加一个部门，比如法务部，会议的数量将增至25次。你怀着美好的心愿——只想让公司继续扩张、交流更加顺畅，最后却创造了一个官僚主义的泥潭。

你发现了一些关于会议爆发的有趣结论：在各个会议上作出的决策彼此之间都有关联，因而，组织某个部分发生的小小变化就会导致遍及整个网络的大规模变化。你的计划导致市场部需要更多的预算，这个预算结果是财务部从客户服务部得来的，客户服务部会说你得让产品的支持难度降低，你将此建议纳入计划，然后循环又回到了市场部需要作出更多调整，循环往复。同样，网络中某个部分的延迟将会导致大规模的堵塞。比如，法务部因审查而暂停了你的产品开发计划，这就意味着市场部无法做财务预算，财务部因此无法将他们的预算告诉客户服务部，然后客户服务部就不知道他们应该聘请多少员工，等等。

此类网络中的相依性造成了考夫曼所谓的复杂性灾难。这种效应的产生是因为随着网络的不断扩展，相依性的数量也会增加，网络某个部

分的积极变化导致其他部分出现消极变化，从而导致发生灾难的可能性也会随着节点数量的增加呈现指数级增长，这又使得发展中的密集连接网络的适应性变得越来越低。

复杂性灾难可以解释为何官僚主义像蒲草一般坚韧。许多公司都开展过消除官僚主义的运动，结果发现几年后春风吹又生。没有谁会故意设计官僚作风，相反，官僚主义是在人们试图实现网络局部最优化时滋生的：财务部是为了确保账务清晰，法务部是为了确保合法有效，而市场部的初衷是推广品牌。问题不在于有人蠢钝或存有坏心，而是网络的发展催生了相依性，相依性导致了冲突约束，而冲突约束使得决策速度变慢，并最终导致了官僚僵局。

可能性VS．自由

如此一来，组织中就存在两股相互矛盾的力量：由节点增加而产生的信息化规模经济以及由冲突限制的逐步发展而导致的规模不经济。这些相互对抗的力量合在一起，有助于我们理解为何大即是美，又为何大即是糟：随着组织规模的增长，其可能性呈指数级增长，与此同时，其自由度呈指数级坍塌。

简而言之，大型组织天然地比其小的时候拥有更多宝贵的机会。理论上来说，大型组织可以做小型组织所能做的一切，有时甚至更多。但想要抓住这些未来的机会必然有得有失，并且这些机会与组织网络结合得越紧密，所要作出的取舍就越痛苦。一些组织的政治理念是这样的：部分群体或部门的局部痛苦足以妨碍组织进入新的状态，哪怕这种状态是更适合全球化的。

IBM公司曾在数十年里利用其体量和规模来主导计算机产业，在整个20世纪80年代，它拥有世界上个人电脑业务的最大份额。[23]然而，1984年，19岁的迈克尔·戴尔（Michael Dell）利用从青少年邮票交易中节省出来的1 000美元创办了一家公司，这家公司在13年后夺走了IBM公司作为全球最大个人电脑销售商的光环，并最终导致IBM公司退出个人电脑市场，将其业务卖给了联想公司。

这样一家世界上最伟大的、拥有数十亿美元资产、在全球拥有数十万优秀雇员、拥有获得诺贝尔奖技术的公司，怎么会输给一个通过集邮换取零花钱的青年呢？我们可以断言，20世纪90年代初，当戴尔公司开始吞食IBM公司的个人电脑市场份额时，IBM公司的一些聪明人肯定会

说："客户似乎喜欢通过直邮的方式购买电脑，戴尔发展很快——为何我们不通过直邮销售电脑呢？"通过直邮销售电脑一定没有超出IBM公司的能力范围，他们完全可以像别人一样，购买一些箱子和气泡包装膜，然后将电脑送去邮寄。为何IBM公司要等到几年后，在戴尔已经超过其市场份额、开始直接向客户销售电脑的时候才行动？原因在于IBM公司陷入了复杂性灾难的陷阱里。

在戴尔公司的起步阶段，IBM公司通过零售渠道向消费者、通过市场人员向企业客户销售了大量个人电脑。如果IBM公司开始通过直邮进行销售，它将招致零售商和销售人员的反对，而这或许会导致销售量即刻下降。此外，毫无疑问的是，有人会质疑这种廉价而新潮的直邮方式是否适合IBM公司这样的一流品牌。事实上，这样的变化将会导致IBM公司的整个业务系统产生连锁反应，从制造工厂到客户服务都会受影响。尽管IBM公司的高管们或许在某些时刻意识到了戴尔是一个威胁，并且他们肯定希望夺回戴尔侵占的市场份额，但IBM公司业务系统的相依性意味着人们在很多时候不得不说不。如果完成一些事情所需要的互动越多，那么出现冲突或限制的可能性就越大。大多数机构中的现实情况是，如果有足够多的人说不，那么事情就不会发生。在IBM公司与戴尔竞争的早期，IBM公司拥有的可能性远远高于戴尔。比如，IBM公司利用直购模式打入企业市场的速度要比戴尔快得多，但公司年龄越大，利用这些机会的自由度就越低。[24]

相依性与适应性之间的张力是网络的一道浓重色彩，深刻地影响着各种类型的系统。软件设计师见它是在程序变得非常复杂，以至于任何升级或漏洞修复都会招致新的漏洞之时；建筑师见它是在客户要将墙移动不过30厘米，但会导致连锁反应，使得项目的成本剧增之时。某些生物学家相信这种张力为有机体的复杂性设置了上限，例如斯图尔特·考夫曼。[25]在经济组织中，规模所带来的益处与复杂性所导致的协调成本和限制之间存在着一种清晰的得失关系。下一个问题是，对此我们能做些什么？

层级的好处

网络理论表明，组织可以采取两种行动，其一是降低连接关系的密度，其二是提高决策的可预见性。我们将分别简要地进行说明。

我们已经讨论了每个节点的连接模式完全相同的网络，比如每个节

点有三个连接点。但如果我们建个层级网络又会怎样呢？对此可以想象有三个员工节点要向一个经理节点汇报，三个经理节点向一个高管节点汇报，以此类推。你想会出现什么情况呢？可以肯定的是，网络的样式会与之前不同，成为密集局部和分散局部的混合体（见图7-2）。考夫曼提出，增加适应性、避免冲突限制的一个有效方式在于将一切打碎。[26] 将层级网络会降低连接关系的密度，从而减弱网络中的相依性。在规模不经济的概念出现之前，层级是扩展网络覆盖面的关键因素，这就是在自然世界和计算机世界里许多网络被嵌在许多网络中的原因。

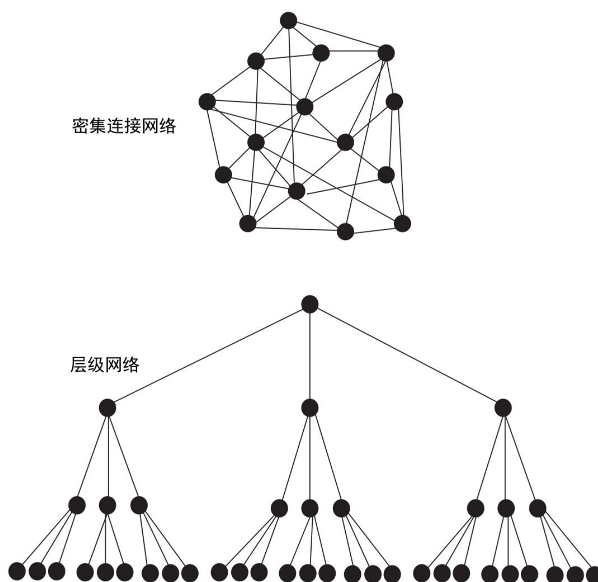


图7-2　密集连接网络与层级网络

在组织环境中，传统观点认为层级是官僚的特征，是适应性降低的原因。经理们被告知，他们应该去除层级，实现组织的扁平化。但与人们的直觉恰恰相反的是，在僵局形成之前，层级可以降低相依性、使得组织覆盖更大范围，从而提高适应性。[27]我们可以用一个简单的案例加以说明。图7-2中的层级结构有40个节点，假设每个节点代表一个人，而组织必须进行一次“让所有人都参与进来”的决策。这个组织中没有层级，那么每个人都必须与其他所有人会面，也就是说总计会发生1　600次会面。但在图中的层级结构中，我们可以假定，每个团队都会与他们的领导会面，获得统一意见，然后将决策推进到下一级别，下一级别的人们又通过会面将决策推进到再下一个级别，以此类推。

如此一来，在底层进行9次会面，在中间层进行3次会面，最后与首

席执行官进行1次会面，总共需要13次会面就可以解决问题，而不是1 600次。谁能想到层级其实能节省会面次数呢？当然，层级有其本身的弊端。比如，信息在沿着链条传递的过程中会有损耗，顶层的人可能会远离一线，糟糕的执行者则会造成很大的破坏。但是，假定层级一无是处确实是一种过分简单化的做法，并且这种做法忽视了层次作为相依性的破除器这一关键定位。

有一种做法是赋予层级结构中的各单元以更多自主性。这是艾尔弗雷德·斯隆（Alfred P. Sloan）的伟大洞见之一，他发明了自治区的概念，使得通用汽车获得发展并成了当时世界上规模最大的公司。斯隆在一家汽车公司内部创建了5家汽车公司，每一家都有自己的品牌并拥有较高的自主性。20世纪80年代和90年代，许多公司开展了为业务单元赋予更多自主性且自负盈亏的举措，这在很大程度上是为了应对组织规模扩大所带来的复杂性。

拆分部门或拆散公司，赋予了这些单元最大限度的自主权。当私募股权公司买下大公司的分部，其业务表现往往会实现巨大的提升。其中大部分原因要归结为给予管理层激励和股权，但在某种程度上可能是因为限制的消除以及自由度的提升。但我们也经常看到相反的案例，即高绩效的小公司被大公司收购后，被大公司的限制压抑到窒息。

无聊更好

考夫曼及其同事在开展原创性工作时发现，平均每个节点拥有一个或两个连接关系的非层级网络展现了自发秩序，但每个节点拥有4个或更多连接关系时就会出现混乱（导致出现大量变化以及复杂性灾难）。[28]随后，来自巴黎高等师范学院的物理学家伯纳德·德里达（Bernard Derrida）与杰拉德·韦斯布奇（Gerard Weisbuch）发现了一个可以改变相变时间点的参数，[29]他们将此参数称为偏差。

在布尔圣诞树串灯案例里的规则是：从周围灯泡得到的输入信号将会转化成输出信号。例如，如果A与B都关闭，那么C也将关闭。现在，假设我们不知道任何灯泡变化的规则——每一个灯泡都是一个黑匣子，那么我们可以通过向单个灯泡输入信号并观察其输出来研究其行为。选择一个灯泡，然后随机输入多个1和多个0，我们的输入流就会有将近50%的多个1和接近50%的多个0。如果输出流也是50%的1和50%的0，那么我们可以说输出是无偏差的；如果输出有90%的1，那么我们可以

说输出偏向1；如果90%的输出都是0，则输出偏向0。我们知道，灯泡的输出是通过对输入应用一种确定性规则计算出来的，因此低偏差并不意味着灯泡的行为是随机的，这只不过意味着其中的神秘决策规则是同样输出1或0。在一位不知道规则的旁观者看来，低偏差节点的行为很难预测，而高偏差的节点更容易预测。

德里达与韦斯布奇发现，在转变成混乱状态之前，偏差越高，网络的连接越密集。如果平均偏差是50：50，那么每个节点的连接数量为2～4个时就会向混乱转变，这跟考夫曼的研究是一致的。如果平均偏差接近75%，那么每个节点的连接数量超过4个才会发生转变。在偏差更高的情况中，网络中节点的连接数量多达6个时才会发生转变。关键在于，节点的行为越有规律，网络所能容纳的连接密度就越大。

在组织中，我们可以将偏差视作可预测性的衡量手段。如果组织的决策是具有可预测性的，就好比灯泡的规则，那么拥有密集连接网络的组织就可以高效运转。然而，如果决策的可预测性相对较低，那么网络中的连接数量就不那么密集，层级更多，并且所需的控制范围就更小。举个例子来说，在部队中，规律的、可预测的行为是受到高度重视的，但在更大的部门中，比如广告创意机构，这种行为就是要避免的。这同样意味着导致行为可预测性降低的因素（例如办公室政治和情绪）会限制组织规模的发展，以至于组织无法被复杂性主导。我们甚至可以据此开出一张创建一个功能失调的组织的配方：将不可预测的行为、扁平的层级以及大量密集的连接混合在一起，这样能够做成事情的可能性就几乎等于零。

秩序的边缘

考夫曼以及其他人的工作得出了一些违反直觉的洞见。IBM公司在应对戴尔公司时出现的问题并不在于一流的公司对变化不敏感，而在于它对变化过于敏感。IBM公司业务系统中密集的相依性以及纠缠不清的互动意味着小小的变动（“我们通过直邮卖电脑吧”）会升级为大问题（“我们有1 000个不能通过直邮卖电脑的理由”）。

这展示了仅在隐喻的层面对复杂性理论进行解读所具有的危险性。许多畅销的管理类书籍和文章都写过关于“混乱的边缘”的内容。这是秩序与混乱之间的边界，有人说大自然的适应性是最强的，对于这种观念存在一种普遍解读——公司之所以不能适应，是因为它们在秩序化的管

理体制中陷得太深，因此需要允许组织有更多的混乱，这样才能进一步激发创新。尽管这种观点听起来很有吸引力，然而，对科学的正确解读实际上应该更细致，并且将会得出不同的结论。

为了理解以上内容，我们可以回到考夫曼设计的闪烁灯泡网络。在考夫曼的研究中，如果平均每只灯泡只与其他两只灯泡相连，那么该网络的行为是十分有序的——小变化不会引起灯泡闪烁模式的大变动。但当他把每只灯泡与其他四只灯泡相连时，行为就会变得很不一样。网络中某个部分的一个小变化（比如稍微改变控制灯泡行为的其中一条开关规则）就会引起重大变化，使得预测灯泡变化的模式变得不可能。在组织中，正是这些重大的变化在后来进化成了冲突约束。[30]每只灯泡从2个连接变成4个连接，就会出现突然的相变，网络面对变化时的精准和不敏感就会变成面对变化时的混乱和过度敏感。如果进化系统对变化的敏感程度处于中等水平，那么系统就能达到最佳的工作状态，其中的原因我们将在后续进行说明。如果进化系统对变化不够敏感，那么这个系统就无法跟上环境变化的步伐。如果系统对于变化过于敏感，那么小变化就会产生大影响。过度敏感的问题在于，如果系统在过去已经取得了成功，那么很少能有重大改变再次提升它成功的可能性，绝大多数的重大变化反而会损害其原本的成功。

考夫曼发现，如果布尔网络中平均每只灯泡拥有2～4个连接，那么系统就会变得具有极强的适应性，进入中间状态。在此状态中，系统通常是有序的，它拥有大型的岛状结构，但在结构边缘也有充满活力的、缓慢渗透的无序状态。系统开关控制规则的微小改变一般会导致结果的小变化，但偶尔，小变化也会产生大影响，有时候这会拖累系统的表现，但有时候也会带来改善。尽管这种特定的网络适应性很强，但考夫曼依然对此很困惑，因为根据大部分自然界或人类组织网络的标准来看，每个节点拥有2～4个连接依然太稀疏了。

如果我们将考夫曼的原始结果与后来的层级和偏差结果相结合，相变的范围就变成了6～9个节点。对布尔网络进行分析所得出的数值与我们通常所见的高效率的人类组织的规模十分接近。[31]比如，在主席或首席执行官领导下的公司董事会或高管委员会通常有5～8位成员。美国最高法院有8位大法官和1位首席大法官，而欧盟的执行委员会有5位副主席和1位主席。一些人类学家推测，这些典型的群体结构和规模源自人类作为狩猎采集者的长期进化传统，而且这样的群体规模是为了组成高效的狩猎群组。久而久之，进化往往非常有效，可以在权衡中找到平衡

点。这些典型的工作群组规模的进化很可能是因为它代表着规模效益（狩猎小组每消耗1卡路里所能获得的食物多于个人每消耗1卡路里所能获得的食物）与复杂性不经济之间的平衡。如果30人的狩猎小组要花几个小时的时间来争论当天是打野牛还是打羚羊，我们的祖先就不可能存活下来。

08

视角四：经济是一种涌现现象

大约在1315年，英国的经济进入了自由落体状态。[1]由于气候条件不好导致连续两季歉收，一蒲式耳[(13)]小麦的价格猛涨了8倍，从5先令涨到了40先令。小麦价格的上涨意味着磨坊主买不起谷物，面包师买不起面粉。农民无法养活他们的牲畜，牲畜随后渐渐患病死亡。牲畜的减少意味着田地无法被耕种，这反过来造成了粮食的进一步短缺。随着粮食供应的崩溃，农民们开始变得绝望，开始吃猫、老鼠、昆虫、树木以及任何能找到的东西。当农民们挨饿时，经济崩溃蔓延到商业领域，木匠和裁缝也陷入了困境。贵族阶层也未能幸免，就连爱德华二世（King Edward II）的宫廷也难以获得食物。英国经济的下滑很快开始蔓延到欧洲大陆，从巴黎到乌得勒支都开始出现萧条。

萧条、衰退和通货膨胀并不仅仅是现代独有的现象，自有历史记录以来它们就在反复出现。[2]经济学中还有其他同样古老的模式，包括第1章讨论的人均财富的长期增长，以及第4章讨论的财富分配。这些模式之所以如此古老，是因为它们必然是经济运作深层原因的结果，而且这些原因是独立于特定时代的技术、政府政策或商业实践的。

在查看经济时序数据的图表时，无论是产出、失业率，还是通货膨胀，人们都会注意到数据非常嘈杂——这些数据曲线始终在晃动。不过，尽管存在波动，数据中还是有规律可循的。例如，图8-1显示了由布兰迪斯大学的大卫·哈克特·费舍尔（David Hackett Fisher）创建的一个指数，该指数提供了从1201年到1993年英格兰近800年的物价通胀情况，我们可以清楚地看到数据中的多个模式。第一是长期的增长趋势；第二是这一增长的顶部是三个驼峰，每一个都持续了几百年，然后是许多较短的起落周期，每一个只持续了短短几年。图中的数据既不是完全规律性的，也不是完全随机的。

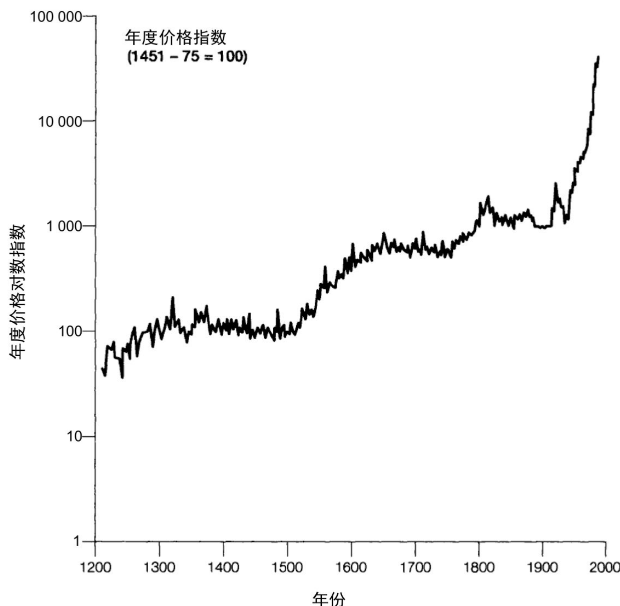


图8-1　1201—1993年英国消费品价格

资料来源：费舍尔（1996年）。

其他的时序数据也具有类似的不太规则、不太随机的特征。例如，美国商业周期的起起落落表现出明显的振荡模式，但振荡的频率从18个月到9年不等，振幅也从轻微的减速到深度衰退不等。[3]经济学家试图利用经济数据中的非规律性历史模式来预测未来的经济行为，但是他们几乎没有成功过。因此，我们有两个问题需要用经济学来解释。首先，为什么经济数据中的模式对经济学家来说不是很有规律，也不是很随机？其次，经济的深层结构特征究竟是什么？

传统经济学历来都难以回答这些问题。如前所述，经济学分为两个部分：微观经济学和宏观经济学。微观经济学是一种自下而上的经济观，它从个体决策者开始，然后逐步建立起市场和经济；宏观经济学是一种自上而下的经济观，它从诸如为什么会出现失业等问题开始，然后深入研究以找到答案。到目前为止，我们主要关注的是微观经济问题，但在这一章中，我们将上升到宏观层面。大多数经济学家都认为，理想情况下不应该有单独的微观经济学或宏观经济学。一个人应该能够从微观行为开始向上归纳，或者从宏观模式开始向下深入，并且能够在一个理论中灵活使用这两种方法。虽然这两种类型中有许多共同的观念、技巧和传统的均衡框架，但不幸的是，我们目前尚未实现这一愿望。就像19世纪建造横贯美国大陆铁路的两个团队一样，微观经济学家和宏观经

济学家一直在这一领域的不同方面努力。但在铺设了一个世纪的铁轨后，他们也未能在中间相遇。我们将简要地看一下微观经济学和宏观经济学对经济模式之谜有什么看法，然后再从复杂经济学的角度来分析这些问题。

商业周期就像抖动的果冻一样吗

商业周期振荡等现象对传统的微观经济学理论提出了挑战。正如前面所讨论的，新古典主义均衡经济学模型并不会自动地翩翩起舞、摆动或振荡。这个模型必须在受到外力冲击的情况下才能活跃起来。这样，均衡模型就充当了传播机制。[4]我们在模型中输入一个震动，它就会传播这个震动，然后产生一个输出模式。想象一下，盘子里有一大堆葡萄果冻，而你用勺子随意敲击果冻堆会发生什么？可以看出，振动可以通过果冻进行传播，从另一边出来，导致果冻发生抖动。我们可以把这些敲击看作外源性的输入，把另一边的抖动看作输出。在传统的经济学文献中，以这种方式工作的模型被称为真实的商业周期模型。这些模型是由芬恩·基德兰德（Fynn Kydland）和爱德华·普雷斯科特（Edward Prescott）在20世纪80年代首创的。[5]

然而，传统的抖动果冻模型存在一个问题。传统的理论通常会将外源性输入描述为随机的，或者说至少是没有可以预测的模式的。但是，如果果冻的输入是随机的，那么输出也将是随机的。信号可以通过传播机制以某种方式进行转换，但输出仍然是随机的。想象一下，你画出了勺击输入和果冻的抖动输出的图形。输出的抖动与输入的敲击不完全匹配，因为果冻会增加敲击的延时，也可能会增加振荡，即一次轻击可产生三到四次抖动。然而，尽管抖动输出看起来与敲击输入不同，但它仍然是随机的。果冻本身是不能接受一个随机的、无序的输入流，并添加复杂的顺序的。[6]果冻是一个均衡系统，如果你停止敲击，它最终会停止抖动。要想让果冻产生真正有序的摆动，唯一的方法就是有序地敲击输入。如果你在果冻上按照摩尔斯电码中的“SOS”（三次快速点击，三次长点击，三次快速点击）敲击，输出的波形将不再是随机的，而是具有SOS信号的结构特点。

微观经济学的真实商业周期模型也存在着一个问题。人们可以输入随机数据，在这种情况下，输出看起来不太符合真实的状况。或者，我们也可以输入具有结构的数据，然后添加一些效果，使输出具有真实数

据不太规则、不太随机的特性。但在后一种情况下，人们并没有真正解释商业周期出现的原因，只是把解释转向了外在因素。[7]

现在我们都是（新）凯恩斯主义者

而从宏观经济学家一直在努力的另一边，我们会看到一些不同的方法。他们必须从数据入手，然后给出一个可能解释。纵观宏观经济学理论的发展历史，宏观经济学家能够令人信服地解释经济的非规则性、非随机性的唯一方法就是背离传统微观经济学的正统理论。因此，他们不得不在不同的方向上铺设铁轨。

当凯恩斯1936年写出《就业、利息和货币通论》时，这种偏离理性均衡的趋势或许发生了一个重要的转变。20世纪30年代，凯恩斯目睹了现代历史上最悲惨的失衡事件之一——经济大萧条。[8]由于新古典主义微观经济学理论认为处于均衡状态的经济是充分就业的经济，因此，这显然必须有一个说法来解释经济为何会如此显著地脱离均衡，以及如何才能加速它的回归。凯恩斯提出的说法，即他的《就业、利息和货币通论》，就是一个动态的理论。[9]

经济世界中的人们会由于某种原因感到焦虑，例如政治不确定性、自然灾害或战争。这种情况下，消费者和商人会变得更加保守，减少开支，同时持有更多的现金。由于经济中任何一个时间点都有一个固定的现金流，所以这些行为会使得货币退出流通循环。这意味着农民、制造商、店主和其他生产者的收入会减少，然后他们会变得紧张起来，进而减少自己的支出和投资，这将导致别人那里的现金也进一步减少。最终，消费和投资的下降导致人们失去了工作，消费支出减少，最终造成了消费和投资下降、失业率上升以及消费者和商人产生焦虑情绪的加速螺旋式下降。

均衡经济学认为，货币供应的这种收缩最终应该是能够自我修正的——价格和工资下降并反映出流通的货币量减少，最终一切都应该回归到正常的充分就业均衡。在大萧条时期，物价和工资确实会下降，但是这种通货紧缩将导致人们花得更少（在通货紧缩的环境下，你的钱在未来会更值钱，所以最好还是持有它），从而进一步加剧恶性循环。凯恩斯认为，这些动态发展可能会导致经济在很长一段时间内陷入失衡状态。因此，为了恢复充分就业，凯恩斯主张政府通过向金融体系注入资金来发挥作用。这种资金的注入会导致支出停止下降，失业率停止上

升，消费者和企业家信心回升，从而将恶性的下降循环逆转为良性的上升循环。

凯恩斯的观点在第二次世界大战后数年内被西方政府广泛接受，同时也在经济学领域引发了长达几十年的纷争。这场辩论在20世纪60年代和70年代达到了高潮，当时米尔顿·弗里德曼辩称，凯恩斯提倡的那种模式的政府支出不会带来长期增长，而只会导致更严重的通货膨胀。在20世纪70年代的高通胀、低增长时期，弗里德曼的观点显得尤其尖锐。弗里德曼在芝加哥大学的同事罗伯特·卢卡斯提出，如果人是完全理性的，那么他们就会明白发生了什么，人们发现经济处于恶性的循环中时就会相应地调整自己的行为，从而使经济回到充分就业的均衡状态。此外，完全理性的消费者和生产者会识破政府的任何干预企图，预见到政府的行动，并以否定政策效果的方式来采取行动。政府干预不仅无法阻止衰退，而且可能会让事情变得更糟糕。卢卡斯的理论在数学角度是非常出色的，他还在1995年获得了诺贝尔奖，但他对理性的超强诠释超出了许多传统经济学家的信任范围。

在卢卡斯之后，加州大学伯克利分校教授、2001年诺贝尔奖的共同获得者乔治·阿克洛夫基于赫伯特·西蒙的观点表明，当人们在超市里购买西红柿的时候，他们是没有理性去估计未来的政府预算赤字的。收集和处理这些信息是不值得的，因为成本太高，太费时间。尽管阿克洛夫没有构建出我们早先看到的归纳理性模型，但他明确表明，消费者和生产者并不完全理性，而这足以引发凯恩斯动力学，使经济陷入衰退。[10]阿克洛夫还承认，时间延迟对经济的动态发展、价格和工资的“黏性”以及它们无法立即自我调整方面发挥着重要的作用。他的模型显示出，在对抗这种恶性循环方面政府可以通过增加市场流动性来发挥建设性的作用。阿克洛夫与埃德蒙·菲尔普斯（Edmund Phelps）、奥利维尔·布兰查德（Olivier Blanchard）和格雷戈里·曼昆等人的作品综合起来，形成了人们现在所说的新凯恩斯主义经济学。[11]虽然新凯恩斯主义经济学在经济学理论家中仍存在一些争议，但是在政府和华尔街的现实世界中，人们普遍认为，政府对利率和预算赤字等因素的调控确实会对经济的表现产生影响。

21世纪初，传统经济学为我们提供了两个相互竞争的假设来解释我们在经济中看到的振荡模式。[12]一方面，我们有基于微观经济提出的真

实商业周期理论，该理论坚持理性均衡观点，认为经济只是在传播外部的冲击影响。这一理论认为，经济振荡的主要原因有外部的政治事件、技术变化等。但是，这些模型不能告诉我们，尽管作为原因的外源性因素发生了巨大变化，但商业周期何以在历史上持续如此之久。另一方面，我们有基于宏观经济的新凯恩斯主义。为了找到内源性的解释，新凯恩斯主义已经背离了传统的经济学正统观点，融入了不完美的理性、动力学和时间延迟性。新凯恩斯主义在很多方面都是向复杂经济学方向迈出的一步，但是，新凯恩斯主义者还没有准备好放弃均衡论，因此我们可以说，该理论的经验成功受到了限制。[13]

“多即不同”

我们之前讨论过复杂适应系统中个体的微观交互如何创建宏观结构和模式。例如，即使是非常简单的糖域因素之间的相互作用，也可能会导致经济增长和收入不平等等结果。复杂经济学的最终成就是发展出一种理论，将我们从个体理论、网络理论和进化理论带到在现实世界看到的宏观模式中。这样一个全面的理论还不存在，但是，我们可以看到它可能是什么样子的曙光。

这一理论将宏观经济模式视为涌现现象，即系统作为一个整体的特征是内生于个体及环境的相互作用中的。[14]我们在前几章已经简要地提到了涌现，但这个概念对你而言可能仍然是神秘的。有什么东西的功能比它各部分功能的总和还要大呢？或者就像物理学家菲利普·安德森说的那样：“为什么‘多即不同’呢？”[15]

涌现可能看起来很神秘，但实际上是我们每天都在经历的事情。假设你能感觉到一个单一的分子，一个由两个氢原子和一个氧原子组成的水分子不会让你感到潮湿，但是杯子里的几十亿个水分子却会让你感觉杯子是湿的。这是因为湿度是水分子在特定温度范围内相互作用的一种集体属性。如果我们降低水的温度，这些分子就会以不同的方式相互作用形成冰，失去其湿润的特性，呈现坚硬的特性。同样，交响曲是一种声音模式，它是从多个乐器的演奏中涌现出来的，而肾脏是一种细胞共同工作的模式，它们一起提供了一种更高层次的功能，这是任何细胞都无法单独完成的。[16]

复杂经济学同样将商业周期、增长和通货膨胀等经济模式视为系统内相互作用产生的涌现现象。复杂适应系统往往具有许多类型系统中常

见的典型涌现模式，这些模式可以帮助我们更好地理解这些系统的工作方式，接下来，我们将研究三种这样的典型模式：振荡、间断均衡和幂律。

振荡：啤酒世界的繁荣与萧条

钟摆左右摆动、吉他弦的振动和心跳都是振荡系统的例子。正如我们已经讨论过的，经济也会随着整个经济的商业周期，以及在第5章中讨论过的行业层次的商品周期的起起落落而振荡，或许还会伴随着更长期的变化浪潮发生振荡。[17]那么，为什么这些振荡会存在，为什么它们在历史上如此持久呢？

振荡是复杂适应系统的一个共同特征。例如，生物生态系统中的族群就会经历振荡。20世纪初，乌克兰化学家阿尔弗雷德·洛特卡（Alfred Lotka）和意大利数学家维托·沃尔泰拉（Vito Volterra）建立了一个著名的模型，用来描述捕食者和猎物之间相互作用所产生的振荡。[18]假设有一群狐狸和兔子，洛特卡-沃尔泰拉模型显示了随着兔子数量的增长，狐狸会吃掉更多的兔子，导致狐狸数量上升、兔子数量下降。兔子数量的减少反过来又导致每只狐狸所能获得的食物减少，最终狐狸的数量也会下降，从而使兔子的数量再次上升，造成狐狸和兔子数量的波动。这个动态系统从不会达到稳定，而是处于无限振荡中。洛特卡和沃尔泰拉的模型中不存在任何外部的冲击驱动着振荡。起伏来自系统本身的结构，而不是任何外部的因素。

那么，经济体系的结构如何产生内生的振荡呢？事实上，让经济系统产生内生振荡并不是很困难：需要的仅仅是一些本科生和几箱想象中的啤酒。20世纪50年代，麻省理工学院的杰伊·福里斯特（Jay Forrester）发明了一款名为“啤酒分销”的游戏，该游戏演示了人类行为与动态结构的结合如何在一个简单的经济系统中相互作用，从而产生振荡。[19]

4个学生志愿者被要求玩一个模拟商品生产和销售的游戏，为了让学生更感兴趣，福里斯特选择了啤酒。4个人中，一个人扮演啤酒酿造者的角色，一个是啤酒经销商，另一个是批发商，还有一个是零售商。酿造者、经销商、批发商和零售商的供应链在许多行业中都是很常见的。在这个游戏中，没有人扮演啤酒消费者，相反，消费者的需求是由一堆面朝下放在零售商旁边的卡片提供的。

在每个回合开始时，零售商从牌堆中翻出一张牌，表示从消费者那里得到了订单（例如4箱），他会向批发商提交订单。批发商查看来自零售商的订单，然后向经销商提交订单，经销商再向啤酒酿造者提交订单。一旦每个人都收到了订单，玩家就通过运送啤酒箱来完成订单。酿酒者把啤酒箱运送给经销商，经销商再运送给批发商，批发商再运送给零售商，零售商再把啤酒卖给消费者。因此，订单在供应链上从消费者流向啤酒酿造者，而啤酒则反向回流。一旦订单提交以后啤酒被运出，下一轮就开始了。

玩家需要为持有啤酒库存支付每箱0.5美元的成本（例如储存啤酒的成本），以及为啤酒耗尽没货可卖支付每箱1美元的成本（例如愤怒的客户和丢失的销售额）。因此，玩家需要保持足够的库存以完成客户的订单而不是耗尽啤酒没货可卖。考虑到成本的不对称，玩家往往更倾向于拥有一些额外的啤酒库存。游戏的赢家是付出代价最少的人。这听起来很简单，但在现实生活中，点啤酒和实际收到啤酒之间存在一定的时间延迟。可以想象，生产啤酒和用卡车装运啤酒都需要一些时间。同样，在提交订单和处理订单之间也有一个小的延迟，你可以将此想象为某人接收订单、在计算机系统中录入订单、进行信用检查等所需的时间。在游戏中，除了啤酒订单，玩家之间不允许有任何沟通和交流。因此，啤酒酿造者不知道零售商那头的客户需求是什么，他们看到的只是来自经销商的订单。

这些时间上的延迟让事情变得有点棘手。假设你是经销商，你从批发商那里接到一大笔订单，这会导致你的库存突然下降，然后你向制造商下了一个大订单来扩充库存，但是这需要几个回合才能收到啤酒，如果同时还有一个大订单呢？你是否应该把订单做得更大些，以期待下一回的持续高需求呢？但如果这只是昙花一现呢？你也可能会在两个回合之后被多余的啤酒淹没。当行动和对这些行动的反应之间存在延迟时，人类就无法表现得很好——当你试图在一个不熟悉的淋浴间调节水温时，你会体会到什么叫“冰火两重天”。

游戏开始时是均衡的，每个玩家都会得到4箱啤酒的订单，交货渠道开始也是满的，所以每个玩家在第一个回合中都会正好收到4箱啤酒，从而保持整个供应链的库存水平不变。从那以后，玩家们要自己来决定下多少订单。玩家不知道的是，代表消费者需求的那套牌的前几张仍然是4箱，但4位参与者的订单却可能比4箱多一点或少一点，这取决于他们愿不愿意冒险。然后，突然间，在某一个回合中，消费者订单牌

从4跳到了8。玩家并不知道这一点，在接下来的游戏时间里，客户订单水平将保持在8。这种订单的增加会给供应链带来干扰。按照传统经济学的观点，这种外部需求带来的冲击只会导致参与者在经过几轮调整后转向新的均衡状态，比如，每个人都下出8箱啤酒的订单，一旦达到新的均衡状态，每个人的库存就会保持不变。

然而，在真人实验中，玩家不可避免地会在库存下降时对需求的激增反应过度，从而过量订购啤酒。随着这波订单过剩的浪潮在供应链上蔓延，它就被放大了。或许这家零售商对需求从4箱增加到8箱感到意外，从而需要补充库存并订购12箱啤酒。然后批发商看到零售商的订单增加到了12箱，就把自己的订单增加到16箱，以此类推。除了反应过度之外，参与者并没有正确地解释下订单和接收啤酒之间的延迟，所以零售商前一个阶段可能订购了12箱，但是还没等到到货库存已经用完了，所以他又订购了12箱。这样造成的结果是，大量啤酒最终开始回流到供应链上，每个人都被啤酒库存淹没了。反应过度的循环会反过来产生以下影响，玩家会大幅削减订单，有些甚至会降至零。因此，订单过剩和订单不足的振荡波会在供应链上下波动，然后我们想象中的啤酒行业就会陷入代价高昂的盛衰周期（见图8-2）。

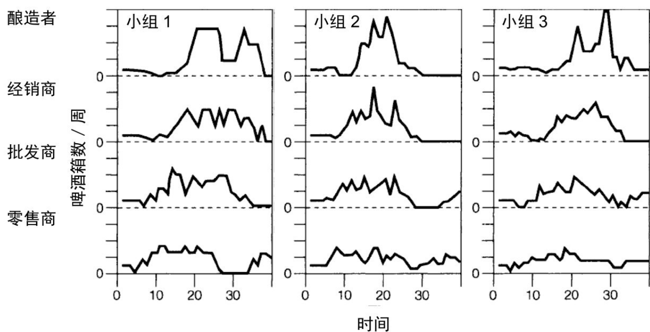


图8-2　啤酒游戏的典型结果

资料来源：斯特曼（2000年）。

麻省理工学院教授约翰·斯特曼和其他人已经在来自世界各地的人中进行了数百次啤酒分销游戏，其中包括工商管理专业的硕士研究生、商人和一些随机挑选的人。[20]他们甚至在专业库存经理和高度理性的经济学家群体中尝试过。然而，结果总是一样，都引发了剧烈的振荡。[21]传统的经济学理论认为，如果参与者是完全理性的（在游戏中，他们拥有完全理性所需要的所有信息），就不应该出现剧烈的振荡，游戏应该整齐地从一个均衡过渡到另一个均衡。如果将实验产生的成本与完全理性

的情况下产生的成本进行比较，那么真人产生的成本平均是完全理性成本的10倍。

究竟是什么样的行为会导致在一个相对简单的环境中出现如此剧烈的振荡呢？斯特曼从统计学上推导出了参与者使用的决策规则。[22]这个规则主要基于心理学文献中被称为锚定和调整的一种行为。参与者不会通过观察游戏板上的筹码库存（他们可以看到的）并结合时间延迟的影响，来演绎计算未来的需要，相反，他们只是查看订单和库存，并归纳性地锚定在看似正常的模式上。因此，他们的“如果-那么”规则试图引导他们遵循正常的模式，参与者可能会将4箱啤酒作为正常的订单模式进行锚定，然后在事情不正常时努力进行调整，例如，“我的库存正在下降，需要订购更多”。在具有时间延迟的环境中，锚定和调整的规则会导致个体过冲和下冲，这反过来又催生了周期性行为的紧急模式。

啤酒分销游戏不仅仅是一种摇摆果冻的传播机制。两者的共同之处在于，游戏受到一个单一的外部冲击的影响——从4箱增加到8箱。但与单一击打中的果冻不同的是，在啤酒游戏中，一旦振动开始，系统就永远不会回到均衡状态。[23]这是因为，啤酒游戏中振荡的最终来源不是外部冲击，而是参与者的行为和系统的反馈结构。系统不传播外部的动力，它是从内部创造了这些驱动力。

这个例子的关键不是简单地把箭头指向完全理性（尽管它做得很好），而是展示变幻莫测的个人行为如何在微观层面上导致宏观层面巨大、紧急的后果，即锚定和调整的规则如何导致了振荡。讲到这里，你可能会问，宏观经济的波动是否是一个比啤酒分销游戏复杂得多的版本的结果。毕竟，经济世界到处都是供应链、库存和时间延迟。宏观经济波动的实际原因无疑是多方面的，但啤酒分销游戏带给我们的启示则是，周期性振荡的原因最终可能在于人们在决策中使用的归纳规则与经济系统的动态结构相互作用的方式。

如果说经济就像一场巨大的啤酒分销游戏，那么其中一个暗示就是，降息和增加政府支出的标准解决方案是无法消除经济周期性振荡的根本原因的：它们只能缓解表象。我们可能永远无法完全消除商业周期性（事实上我们可能也不想完全消除商业周期性，因为周期性收缩不仅清除掉了对资源的低效使用，而且刺激了创新），但是，如果政府想用一种更根本的方式来减弱周期性振荡的影响，那么就需要考虑经济体系本身的结构。

有证据表明，即使没有政府的明确干预，经济的动态结构也总是在

发生变化的。有两种方法可以抑制啤酒分销游戏中的周期性：一种是缩短时间延迟，另一种是给参与者提供更多的信息，例如让啤酒酿造者直接看到在零售层面发生了什么。[24]始于20世纪60年代的信息技术革命直接影响了这两个因素。有数据显示，自1959年以来，美国商业周期的波动性一直在减弱。尤其是随着价格低廉、运算速度快的电脑的广泛使用，自20世纪80年代以来，其波动性开始大幅下降。[25]计算机能够使公司加快订单处理速度，采用准时制库存方法，并通过电子手段把生产者与其供应链联系起来。关于波动性的降低有多少可归因于技术和商业惯例的变化，经济学领域也存在一些争论。[26]但是，很明显，宏观经济的啤酒分销游戏正在发生变化。[27]

间断均衡：有“关键物种”技术吗

在达尔文的《物种起源》出版后的一个世纪里，生物学家都认为进化是以一种庄严而相对线性的方式进行的，从而导致了物种形成和灭绝的平稳模式。后来，古生物学家奈尔斯·埃尔德雷奇（Niles Eldredge）和斯蒂芬·杰伊·古尔德（Stephen Jay Gould）在1972年发表的一篇具有里程碑意义的论文中推翻了这种传统观点，他们认为化石记录表明，生物进化的道路并不是平坦的。[28]相反，它经历了长时间的相对停滞，其间穿插着爆炸性创新时期和大规模灭绝时期。例如，在大约5.5亿年前的寒武纪时期，一场进化创新的爆发见证了多细胞生命接管地球，并创造了今天地球上大部分的生物门类。然后，大约在2.45亿年前的二叠纪晚期，古尔德称之为“所有物种灭绝的始祖”，当时地球上96%的海洋物种都消失了。[29]古尔德创造了“间断均衡”这个术语来描述平静和风暴交替出现的模式（这个术语在技术上并不完全正确，因为从数学意义上讲，生物进化从未真正处于均衡状态），古尔德指出，他想用这个词来表达的是平静或停滞的时期，其间穿插着变化的时期，但“间断均衡”这个短语一直存在，我们也将在本书中使用它。[30]

间断均衡模式不仅出现在生物进化过程中，也出现在从雪崩到股市崩盘的其他复杂系统中。[31]复杂系统研究人员从20世纪80年代就在研究这些模式及其起源，得出的结论之一是，这种模式产生的一个重要因素是系统中相互作用的网络结构。正如我们在第7章中所讨论的，由圣塔菲研究所的邓肯·瓦茨和马克·纽曼所做的研究表明，许多类型的网络自发组织成一种结构，其中混合了非常密集的连接和非常稀疏的连接。[32]印度科学研究所的研究员桑贾伊·贾因（Sanjay Jain）和他在尼赫鲁中心的

同事桑迪普·克里希纳（Sandeep Krishna）认为，正是这种网络结构构成了生物生态系统中间断均衡出现的基础。[33]

贾因和克里希纳创建了一个不断进化的计算机生物生态系统模型。研究人员发现，如果他们随机从模拟的生态系统中移除掉一些“物种”，通常都不会发生什么。然而，移除一个随机的物种偶尔也会引发一连串事件，最后导致大规模物种灭绝。某些物种在食物关系和生态位竞争的网络中与其他物种紧密相连，生物学家称这些物种为关键物种。例如，我们可以想象一种阿米巴变形虫，它是各种昆虫和滤食性蠕虫的食物来源，而这些昆虫和滤食性蠕虫又是各种鸟类和哺乳动物的食物，而这些鸟类和哺乳动物又反过来对各种植物物种产生影响。这样一来，位于食物链底部的阿米巴变形虫数量的突然下降，就可能会在整个生态系统中造成广泛的影响。

在他们的生态系统模型中，贾因和克里希纳指出了间断均衡模式的三个不同阶段。首先，在一个随机的阶段，网络没有太多结构渗透，随机变化也没有太多的效果——这是间断均衡模式中的一个“均衡”时期。其次，一项创新使网络突然进入了增长阶段。触发这些阶段的创新往往会在正反馈循环中催化其他的创新，从而带来进一步的创新。最后，当新物种出现并在生态系统中占据一席之地时，它们就创造了一个食物和生态位网络的秩序。增长阶段最终趋于平缓，随后出现的是一个有组织的阶段。在这一阶段，变化得到巩固，网络高度结构化，关键物种出现在交互网络的关键点上，就像航空公司路线图中的枢纽城市。在有组织的阶段（另一个“均衡”时期），网络会持续泡沫化一段时间，但随后一项创新或随机变化会击中某一个关键物种。影响到某一个关键物种的变化辐射到结构中，然后网络会在一波灭绝中崩溃。随后，该过程再次以新的随机阶段开始，从而带来另一段增长。

包括马克思和约瑟夫·熊彼特在内的许多观察家都注意到，技术创新会以类似的平静和风暴模式进行。[34]大多数技术创新的影响是有限的。例如，在1957年，通用电气公司发明了石英卤素灯。这项技术最终在20世纪80年代变得足够便宜，卤素灯成为一种流行的消费品。卤素灯是对标准白炽灯泡的一个重要改进，尽管如此，人们也不能说它们对社会产生了巨大的影响。

与此形成对比的是另一项发明，它最初也是对前一种技术的渐进式改进。1901年12月，伽利尔摩·马可尼（Guglielmo Marconi）通过无线电波从英国的康沃尔横跨大西洋向纽芬兰发送了三次摩尔斯电码。马可

尼的目的本来就是要找到比有线电报更便宜、更方便的替代品，而当时并没有多少人对他的发明印象深刻。英国《金融时报》援引英美电缆公司的说法对这项新技术表示不屑一顾："无线电报在商业应用上的可能性是如此遥远，以至于我们对未来毫无预警。"[35]剩下的，就像他们说的，就是历史了。无线电报的出现产生了巨大的影响，反过来又催生了电视、微波通信、雷达、移动电话以及现在的无线互联网接入的发明。可以说，马可尼的发明带来了翻天覆地的变化，极大地改变了大众文化、娱乐、政治，甚至是军事战略。

多年来，人们对技术发展作为一个进化过程进行了大量的研究，[36]这也是我们将在第11章中讨论的话题。我们在这里想要说明的是，这项工作的两个关键观察结果都与对间断均衡的讨论有关。首先，任何技术的发展都不是孤立的。所有技术都依赖于与其他技术组成的关系网络，例如移动电话的发明不仅涉及无线电技术，而且涉及许多其他领域，如计算机技术和编码技术。[37]这些相互关系不仅是技术上的，也是经济上的。例如，围绕汽车发展起来的经济网络就包括了钢铁制造、石油、酒店和快餐等行业。[38]

正如哈佛大学商学院的金·克拉克（Kim Clark）所指出的，技术本质上是模块化的，例如汽车由发动机、变速器、车身等组成。[39]人们可以将模块组装成"架构"，在本例中是汽车本身的设计。模块的创新可以形成新的架构，但是架构的创新往往会对创新产生巨大的催化涟漪效应。因此，有两个关键的特征导致了贾因和克里希纳模型中的间断均衡模式——稀疏、密集的相互作用网络，以及来自单个节点的催化效应。我们将在第11章中对此进行更详细的探讨，从中可以看到技术网络如何受到级联变化的影响，从而催生了间断均衡的新兴模式，而某些技术在这些网络中发挥着关键作用。

幂律：地震和股市

在前文中，我曾简要地指出传统经济学的一个预测是，股价应该遵循随机游走模式。这就好比你将笔放在一张纸上，然后从左向右拉笔，在这样做的同时，你还要以随机增量向上或向下移动笔。你也可以想象一个喝醉酒的人在大街上蹒跚而行的情形。图8-3引自耶鲁大学数学家贝努瓦·曼德尔布罗特（Benoit Mandelbrot）的著作。其中一个是IBM公司股票在1959年到1996年间的价格（以对数标度绘制），另一个是随机

游走数据的样本（这个随机游走有增长的趋势）。这些图看起来非常相似，如果没有标记，你很难猜出哪个图是IBM公司的股票价格，哪个是随机游走的数据样本。从中不难看出为什么随机游走在100多年来一直是金融理论的核心部分。

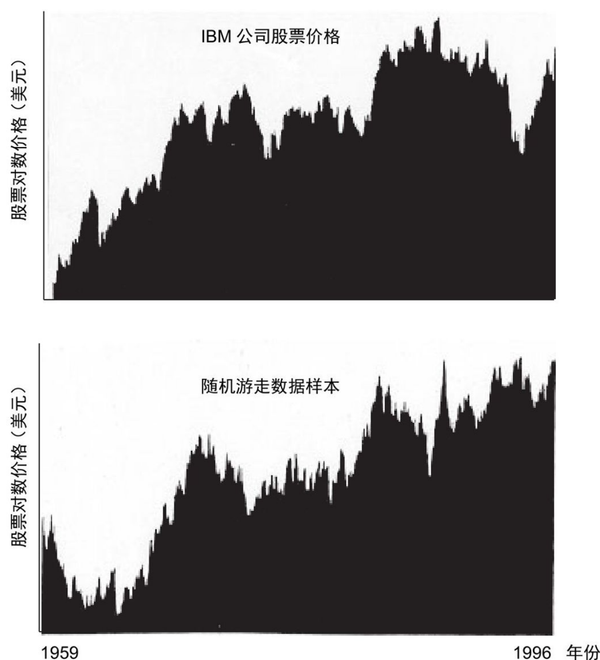


图8-3　IBM公司的股票价格与随机游走数据的样本

资料来源：曼德尔布罗特（1997年）。

在不考虑绝对价格水平的情况下，让我们先来看看股票价格在每一天的涨跌幅度，如图8-4所示。这两个图看起来差别很明显，上面的图非常尖锐，价格波动较大的时期和价格波动较小的时期在时间上是重合的。而下面的图形比较模糊，随着时间的推移，大的移动和小的移动随机混合在一起。从这个角度来看，IBM公司的数据显然不太像模糊的随机游走的数据。正如在第3章中提到的，几位研究人员已经在统计上证明股票并不遵循随机游走规则。IBM公司股票价格变动的块状模式表明，价格的波动性与时间是相关的，这也是我们在上一节讨论的间断均衡的风暴-平静-风暴模式。除此之外，你可能还会注意到一些别的东西，比如随机游走数据图中存在一些较大的价格波动，但没有一次价格波动是真正明显突出的。真实的股票价格数据中往往有几个点要么在上面飙升，要么远远低于其他样本，那么是什么导致了如此剧烈的价格波动呢？

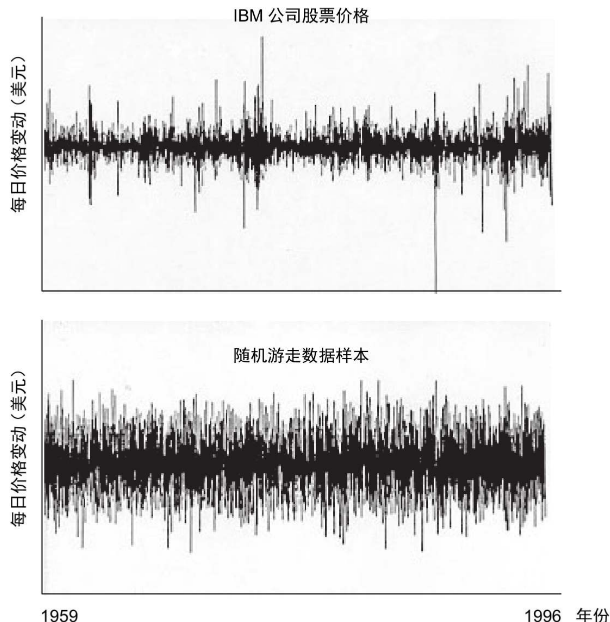


图8-4 股票价格变化图

资料来源：曼德尔布罗特（1997年）。

传统经济学告诉我们，每当有重大消息冲击市场时，股价就会波动。传统理论的预测之一就是，大幅的价格波动应与重大的突发新闻或消息相对应。大卫·卡特勒（David Cutler）、詹姆斯·波特巴（James Poterba）和拉里·萨默斯在1989年进行的一项研究中验证了这一预测。[40]他们观察了1941年至1987年间美国股市的大幅波动，然后浏览了当时的报纸来确定与这些波动相对应的新闻。结果表明，通常情况下，发生严重崩盘的日子里并没什么大新闻可说。如前所述，《纽约时报》对1987年10月19日标准普尔500指数下跌20%的解释是“对美元下跌和贸易赤字的担忧”。1946年9月3日是另一次大崩盘的日子，当时媒体的头条新闻是“价格攻击没有根本原因”。卡特勒和他的同事们研究了当时最大的新闻事件，并测量了当时的市场走势。例如，当日本轰炸珍珠港时，当时的股票市场只下跌了4.4%；当古巴导弹危机得到和平解决后，股票市场只上涨了2.2%。为什么市场上有这么多与新闻无关的剧烈波动？这个谜题的答案存在于一项有趣的观察中：股价走势看上去不太像随机游走，而像另一种现象：地震。

20世纪50年代，来自加州理工学院的两名地球物理学家贝诺·谷登堡（Beno Gutenberg）和查尔斯·克里特（Charles Richter）走进加州理工学院的图书馆，发掘描述地震的资料。[41]他们感兴趣的是每种不同强度

的地震发生了多少次。例如，强烈地震是常见的还是罕见的？2级地震发生的可能性是4级地震的两倍吗？他们将数据分成若干个等级，一个等级为2.0～2.5，下一个等级为2.5～3.0，以此类推。然后他们绘制了数据的分布曲线，类似于糖域模型中的收入数据图。

最著名的数据分布图是钟形曲线，又叫正态分布或高斯分布，弗里德里克·高斯（Karl Friedrich Gauss）在19世纪发现了这种分布。如果我们收集一个群体中女性的身高数据，为每一个身高范围创建一个等级（例如130cm～135cm，135cm～140cm，等等），并计算不同身高的女性出现的次数，我们就会得到一个平滑的钟形曲线。曲线中间的驼峰表明一个典型的女性身高在150cm～175cm之间，而两端则表明很少有非常矮的女性或非常高的女性。

对于地震，人们可能会期望它的等级分布看起来类似于上面的女性身高分布图。我们可以假设地震有一个典型的级别，大多数地震都在这个范围内，很少有小地震或大地震。但是当谷登堡和克里特把找到的数据绘制成图表时，他们发现了一些完全不同的东西。为了使结果更容易看清楚，他们在两个坐标轴上用对数比例尺绘制了数据图，而双对数图上的数据结果几乎是一条直线（见图8-5）。双对数尺度上的这条直线意味着，对于地震分布来说，它的中间没有类似身高分布的“典型”大小。确切地说，地震的震级多有不同，但是震级越高，就越罕见。具体来说，地震能量每增加一倍，相应震级的地震发生的概率就会降低3/4倍，最终人们得到了一个从小规模地震到大规模地震分布的滑坡曲线。物理学家称这种关系为幂律，因为这种分布是用指数方程或幂来描述的。[42]

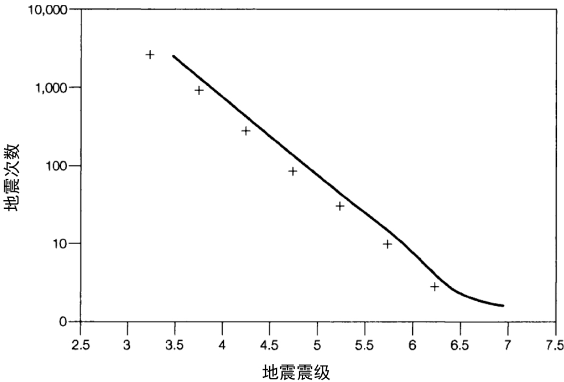


图8-5　地震震级的幂律分布

资料来源：数据来自南加州地震中心，图来自布坎南（Buchanan，2000年）。

人们在各种各样的现象中发现了幂律，包括生物灭绝事件的规模、太阳耀斑的强度、城市规模的排序、交通堵塞、棉花价格、战争死亡人数，甚至社交网络中性伴侣的分布。[43]幂律和振荡、间断均衡一样，是复杂适应系统的另一个特征。

地震并不是第一个被发现的幂律现象，第一个幂律现象是在1895年由帕累托在经济学中发现的，尽管当时它还没有被认为是幂律。[44]帕累托经过对人们收入的研究发现，很多人都是穷人，中产阶级的范围很广，而只有极少数人是超级富豪。人们的收入每增加1%，相应的家庭数量就会减少1.5%——在双对数图上表现出来是一条直线，即幂律。帕累托没有数学工具来充分理解其发现的重要性。[45]20世纪60年代，当贝努瓦·曼德尔布罗特对芝加哥商品交易所棉花价格的波动感兴趣时，幂律在经济学中重新出现了。曼德尔布罗特在绘制数据图表时很快注意到，就像IBM公司股票价格一样，棉花价格的波动性比传统理论预测的要大得多。此外，他还注意到波动似乎没有自然的时间尺度。如果他从图上取一小部分，比如一个小时，再把它拉长到一天的长度，人们根本无法分辨出哪个图代表的是一个小时的数据，哪个是一天的数据。曼德尔布罗特查看了包括黄金和小麦在内的其他商品的数据，也发现了同样的模式——幂律。[46]曼德尔布罗特第一次发现这个理论时，他的工作在很大程度上被经济学家忽略了，一部分的原因是他是一位来自该领域之外的数学家，另一部分的原因是数据与传统理论不相符。

在20世纪80年代和90年代之前，幂律基本上没有得到人们的重视，直到有关经济是一种复杂适应系统的观点被重视才引发了人们对幂律的兴趣。物理学家在分析自然系统的幂律方面经验丰富，一些“经济物理学家”开始研究股市的数据。其中，波士顿大学的吉恩·斯坦利（Gene Stanley）计算出，如果股市像传统经济学假设的那样遵循随机游走规则，那么1987年的黑色星期一发生崩盘的概率是10^{-148}%。宇宙中已知最小的测量单位是普朗克长度，是10^{-33}厘米，在我们的想象中，市场不太可能只是随机地陷入如此严重的崩盘状态。高斯分布、随机游走的波动几乎从来没有超过5个标准差，然而，在真实的经济数据中，比如股市崩盘，5个甚至更大的标准差事件确实会发生。[47]

为了弄清楚究竟发生了什么，在1994到1995年中，斯坦利和他的团队针对美国最大的1 000家公司每隔5分钟进行一次股票市场交易数据的取样，共获取了4 000万个数据点。经过分析发现，股票价格的波动在分布的尾部遵循明显的幂律。为了确认这一点，他们还研究了1961年

至1996年35年间6 000支美国股票的3 000万份日交易记录，这些记录再次形成了幂律。在传统的经济学家看来，或许斯坦利研究的5分钟内的数据是非高斯分布的，但在更长的时间内，数据仍应该是呈高斯分布的。斯坦利还将他的时间周期改变了三个数量级，从5分钟到6 240分钟（16天——除此之外，很难有足够的数据得出有力的结论）。这些数据看起来确实比短周期的数据更接近高斯分布，但是它仍然遵循幂律。[48]

这一结论的后果之一就是，金融市场的波动性远高于传统经济学让我们相信的水平。如果市场遵循幂律，那么黑色星期一事件发生的概率更接近10^{-5}（这意味着在每100年里都有可能发生一次），而不是10^{-148}，这是一个巨大的差异，显然对投资者如何思考和管理风险有重大影响。其他的经济数据也显示出让人吃惊的、清晰的幂律规则。罗伯特·阿克斯特尔利用1997年的美国人口普查数据，对所有拥有一名或多名员工的550万家公司进行了分析，结果显示，以员工数量来衡量的公司规模也符合幂律分布。[49]斯坦利和他的团队发现，公司的销售业绩增长以及国家的GDP增长同样遵循幂律。[50]我们将在后面的章节中探讨这些发现的含义。

为什么股票市场如此不稳定

为什么股票市场比传统理论预测的更不稳定？为什么其波动性会遵循幂律分布呢？多因·法默是一位物理学家，曾参加过圣塔菲研究所最初举办的经济学会议，他和他的团队相信自己已经找到了这些问题的答案。[51]关键在于，大多数的证券交易所可以进行两种类型的交易，一种是市价单，其中交易员会以现在最优惠的价格购买（或出售）股票；一种是限价单，如果价格下跌到100美元，交易员就会购买股票（或者相反，如果价格上涨到100美元则卖出股票）。在本例中，100美元是交易员准备交易的上限。市场中的每一支股票都有一个限价单簿，用于跟踪限价单。过去的账簿是一个分类账，一边是买入指令，另一边是卖出指令，而今天这本账簿是电子版的。

实际上，限价单簿是未成交订单的库存或存储设备。如果交易者将股票的买入限价单定为100美元，而当前价格是110美元，那么该限价单将被存储在账簿中，直到该股票跌至100美元或被取消。在任何时间点上，人们都可以查看限价单簿，查看最佳买入和最佳卖出的报价。例如，最佳买入报价可能是100美元，最佳卖出报价可能是102美元，最佳

买入和卖出报价之间的差距被称为买卖差价。

现在让我们看看当一个新的市价单到达限价单簿时会发生什么。股票交易所关于订单如何完成的具体细节略有不同，纽约交易所还保留着人工交易，而伦敦交易所则完全是电子化的，但大多数交易所都有两条共同的规则：价格优先和时间优先。假设市价单要求以你目前能得到的最优价格供货，价格优先意味着我们要从账簿中的最优价格开始，尽可能多地供应订单，然后转移到下一个最优价格，以此类推。时间优先意味着如果账簿中有两个限价单，且价格相同，则先供应较早的订单。

为了说明这一点，请想象以下场景。你打电话给股票经纪人，要求购买1 000股X股票，他会把你的要求传达给交易所。账簿中目前的最佳销售报价是102美元的限价单，此时在这个价格下有200股股票存在。系统按这个价格供应你的订单中的200股，另有800股待供应。然后系统进一步查找限价单簿，寻找下一个最佳价格，即300股限价卖出订单，价格为105美元。你买了这些股票，订单上还有500股待供应。再往上追溯，下一个最佳价格是200股的限价卖出订单和600股限价卖出订单，价格都是107美元。假设200股在账簿中出现的时间更长，利用时间优先法则，系统会自动分配这些股票，加上从第二部分分配给你的300股，总共是500股。现在你的订单已经完成了，成交均价为105.40美元。考虑到账簿的当前状态，这是买入1 000股的最佳价格。在你的1 000股交易完成后，仍有300股以107美元的价格剩余下来了，所以这是现在新的最佳要价。你的1 000股市价单的影响是将要价从102美元提高到了107美元。正如法默所说，可以想象，随着时间的推移，限价单像雪花一样不断涌入账簿中，并以不同的价格水平累积在账簿中。然后每隔一段时间，市价单（或买卖差价内的限价单）会触及账簿并触发一堆限价单，从而上调或下调股票价格。

法默和他的团队想要了解这个订单履行的过程以及限价单的结构是如何影响股票价格的。他们分析了伦敦证券交易所的一组数据，该数据显示了16种最大的和交易量最高的股票的交易，并且每笔交易的数据都能完整地看到。总的来说，研究人员分析了4 000多万个“事件”（下订单和取消订单）。法默和他的团队发现，造成价格大幅波动的原因是订单簿本身的结构——当订单价格之间存在较大差距时，就会出现较大的波动。以全球制药公司阿斯利康的股票交易为例，在某些特定的时刻，阿斯利康的限价单有一个小的限价卖单——31.84英镑，然后是一个大缺口，之后下一个限价卖单为32.30英镑。随后出现了一个非常小的市场

买单，要价在一次交易中从31.84英镑跃升至32.30英镑，涨幅为46便士（1.4%，见图8-6）。这个价值仅为16 000英镑（28 000美元）的微小交易，将股价提高了23便士（以买卖差价的中点计算，因此就是要价中的46便士涨幅的一半），并让阿斯利康公司的总市值增加了3.74亿英镑（6.58亿美元）。

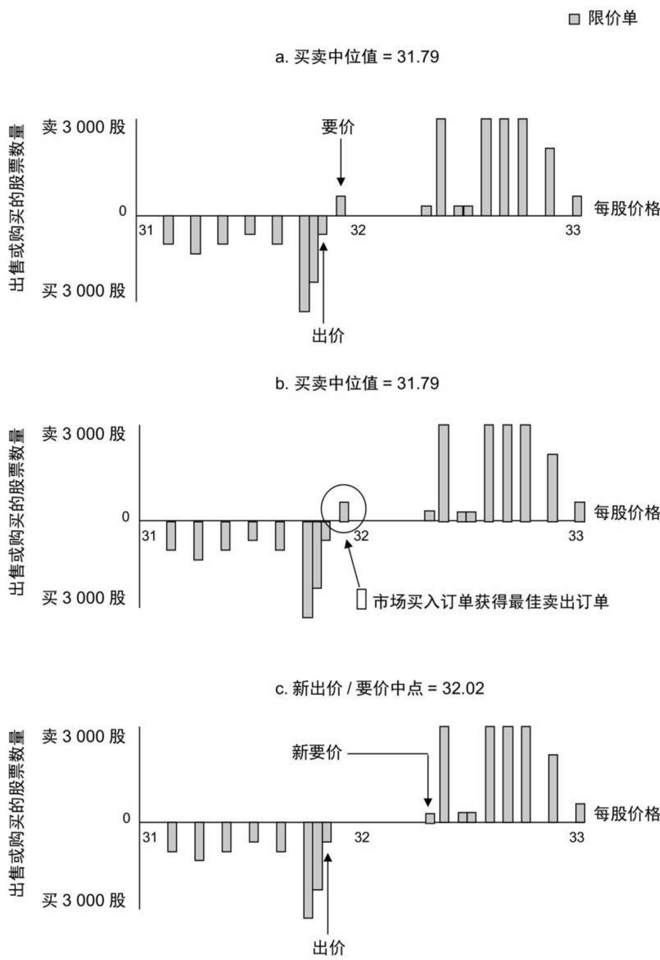


图8-6　阿斯利康限价单中小订单的价格大变动

资料来源：法默等人（2004年）。

然而，那天报纸上并没有什么大新闻，阿斯利康也没有采取任何措施来改变公司的市场价值。市场价值的这种巨大变化仅仅是一种订单模式的人为因素，它随着时间的推移存储在账簿中。在进一步的分析中，法默和团队发现这样的事件相当普遍。即使是规模最大、流动性最强的股票，通常也只有30个左右的价格水平，其中充斥着限价单以及大量不

受欢迎的价格，因此也存在缺口。你可能会认为大多数限价单都会集中在最好的出价和要价附近，这在某种程度上是正确的。但是法默他们也发现，订单分散在整个账簿中。例如，某一个人可能只坚持一个高于当前价格的卖单方式，然后让它留在那里，并希望价格最终会上涨，从而在账簿中创建一个零碎的订单和缺口的模式。

专业的交易员知道，即使是大股票，其流动性也比人们想象得要小，而且订单簿也相当零碎。这就是为什么他们必须在一段时间内进行大量的交易——大宗交易同时冲击市场会使价格波动过大。但直到法默和其同事进行研究后，人们才意识到，这种订单簿上的零散事件会造成多大的股市波动。该团队建立了一个有关订单簿的数学模型并表明，订单簿本身的结构就是波动性的一个重要来源。正如人们可能预期的那样，这些数据还显示，规模较小、交易不太频繁的股票比流动性更强的大股票波动性更大。

在随后的一项研究中，法默和他的同事迈克·斯扎布斯（Mike Szabolcs）进一步研究了限价单簿上的订单放置模式。[52]他们关注的变量是从当前最佳买卖差价到新限价单之间的距离。例如当前的买卖差价是100美元到102美元，那么卖出限价单到达101美元、102美元、103美元等价格的概率是多少？或者买进限价单到达101美元、100美元、99美元等价格的概率是多少？他们发现，订单遵循着一种非常规则的模式，即学生氏分布，固定在买卖价差的中心位置。法默和斯扎布斯指出，订单模式的规律性意味着，下订单的交易员的行为也有一定的规律性——这个结果与传统的所有交易都是由不可预测的新闻事件驱动的理论是不一致的。当法默和斯扎布斯将这个研究结果与之前对限价单的研究结合起来之后，他们非常接近地再现了研究的股票表现出的幂律波动。

这一结果并不是说真正的大新闻对股市不重要。如果阿斯利康公司公布的业绩令投资者感到意外，其股票将作出相应的回应。在许多方面，限价单簿是一种记忆形式，或者说是过去新闻的存储库，因为订单中的订单模式可能在下单时受到了新闻的影响。但结果也确实意味着，股票价格的变动可能与当前的新闻无关，而是一个新的订单流和特定的订单模式之间相互作用的产物，而这些订单恰好在某个时刻存储在订单簿中。一些传统的经济学家可能会试图将这一运动视为可以忽略的短期噪音而不予理睬，而法默的回答是具有双重含义的。首先，限价单簿的影响不仅仅是短期的。他和他的团队成员表明，同样的波动性分布也出现在更长的时间尺度上。其次，它不是随机噪声。价格运动的幂律模式

是从系统本身的结构中产生的，就像啤酒游戏中的振荡是从系统的结构中产生的一样。法默的结论只是针对个别的股票的。在第17章中，我们将探讨“无新闻的波动性”如何与真实的人类股票交易员的行为规则相结合，从而在整个市场上制造出地震的效果。

在很多方面，啤酒分销游戏和法默团队的模式是一样的。商业周期和股价波动等复杂的涌现现象可能有三个根本的原因，首先是系统中参与者的行为。正如我们所看到的，真实的人类具有真实的行为规律性，无论是啤酒分销游戏参与者的锚定和调整法则，还是导致股票订购中产生学生氏分布的尚未被理解的规律性。其次是系统的制度结构有很大的不同。在啤酒分销游戏的例子中，酿造者和零售商之间的供应链结构产生了动态变化，当这与参与者的行为相结合时就会产生波动。在股票市场的例子中，限价单系统的结构与交易员的行为相结合就会导致幂律波动。最后是系统的外源性输入。就啤酒分销游戏而言，主要指客户订单的一次剧增；就股市而言，主要指新闻。这些外源性冲击无疑会引发并推动系统的动态变化。在外生因素发挥作用的同时，传统经济学的均衡论的局限性导致了对这一因素的过分强调，而牺牲了其他两个因素。

复杂经济学并不是经济模式之谜的答案，但它为我们提供了理解这些不同的因素如何结合在一起并导致了我们观察到的行为结果的新工具。现实世界的经济是一个比传统经济学想象的均衡世界有趣得多的地方。不稳定的振荡、间断均衡和幂律，这些都是复杂的适应性经济在起作用时的标志性行为。

09

视角五：
经济是一个进化系统

你有多少次听到过“这是一片丛林”或者“适者生存”这样的话？在谈论经济时，商人、记者和学者都很自然地倾向于使用生态系统和进化这样的词。复杂经济学最有力的主张之一就是，这种语言不仅仅是隐喻——组织、市场和经济不仅仅是像进化系统，它们确实就是进化系统。

在这一章中，我们将看到进化不仅仅是生物学的内容。相反，进化是一种为复杂问题寻找新颖的解决方案的通用而强大的方法，它是一种学习算法，不仅能够适应不断变化的环境，而且随着时间的推移能够积累知识。它是负责自然界所有秩序、复杂性和多样性的准则。正如我们将在第三部分中看到的，进化同样也是经济世界所有秩序、复杂性、多样性以及财富的准则。

没有设计师的设计

丹尼尔·丹尼特是塔夫茨大学认知研究中心的进化理论家和主任，他说进化是一种“无须设计师的创造设计”方法。[1]当我们认为某些东西具有设计时，这说明它适合于一项任务。锤子是用来钉钉子的，细菌是被设计出来在特定环境中生存和繁殖的。我们还认为被设计的事物具有一定的复杂性、顺序和结构。我们不会认为海滩上的沙粒是被设计出来的，但会认为喷气式发动机的复杂结构、鹦鹉螺壳的螺旋室或者一段音乐中复杂的音符排列都是被设计出来的。正是这种适用于目的性和复杂性的结合，将设计的东西与非设计的东西区分开来。露出地面的条状岩石在人们看来可能是很美的，甚至可能看起来像一件艺术品。但它没有功能，不适合于任何特定的目的，是随机地质力量的偶然创造品。设计出来的东西熵比较低，它们绝不是随机产生的。

我们只在两个领域看到了设计：生物和由生物创造的人工产品。在生物学里，袋鼠的腿是用来跳跃的，蝙蝠的声呐接收器是用来在黑暗中

寻找东西的，花的雄蕊巧妙地伪装成蜜蜂的性器官是为了促进授粉。在人类世界里，螺丝刀被用来拧螺丝，抹刀被用来烤薄饼，大型喷气式飞机被用来运送人。但人类并不是唯一可以创造并展示设计物品的生物，白蚁可以精心建构巢穴，而海狸可以建造复杂的水坝。

威廉·佩利（William Paley）是英国圣公会的神父和哲学家，他在1802年出版的著作《自然神学》（*Natural Theology*）中指出，像手表这样复杂和被设计出来的东西都是以钟表匠为前提的。因此他认为，自然界的复杂性和设计性表明一定存在一个神圣的钟表匠。高度设计的东西不会凭空冒出来，毕竟，设计表现出了目的、智慧和解决问题的能力（这与两个世纪后所谓“智慧设计理论”的支持者提出的论点是相同的）。[2]

然而，这正是进化所做的——它自己创造设计。牛津大学进化论学者理查德·道金斯（Riehard Dawkins）的著作也为这一章提供了很多信息，他把进化称为“盲眼钟表匠”。[3]进化是一个无意识的、机械的、简单的准则，同时，它在创造巧妙的设计方面也非常高效。

人工生命

1994年，麻省理工学院媒体实验室前成员、现供职于基因艺术公司（GenArts）的卡尔·西姆斯（Karl Sims）想要研究行为的进化。他想以一种更快的方式来完成这项工作，这样就能得到比在实验室里用细菌和果蝇做实验更多的控制。因此，卡尔·西姆斯在一台超级计算机上创建了一个模拟的进化世界（没有超级计算机和卡尔·西姆斯的编程天赋的人可以通过各种商业电脑游戏来体验、模拟进化，如Maxis的SimLife游戏和微软的Impossible Creatures游戏）。[4]模拟世界中，每个生物的身体都由一系列相互连接的矩形积木组成。生物可以拥有任意数量的躯干积木，躯干积木还可以是任意尺寸的（例如立方体、短粗矩形、细长矩形）。这些积木以铰链的方式相互连接，并且可以弯曲（见图9-1）。通过一个简单的计算机“大脑”协调运动，积木生物可以控制其铰链积木躯干部分的运动。西姆斯给每一个积木生物设定一个目标，然后该生物会用身体上的传感器来评估它所处的位置与目标之间的关系。然后，积木生物会试图采取行动，通过以某种方式移动其铰接块体部件来实现目标。在西姆斯的第一个实验中，这些生物的目标是模拟在水中以最快的速度游泳。

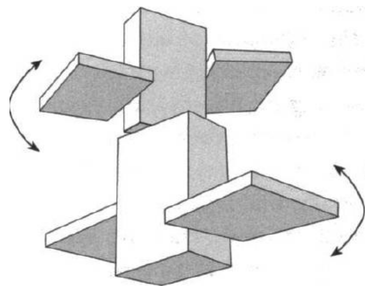


图9-1　卡尔·西姆斯的积木生物

每个生物都由可以移动的积木块组成。

资料来源：西姆斯（1994年）。

西姆斯通过赋予每个积木生物一种计算机DNA，为他的积木世界带来了一个生物学上的转折，这种DNA描述了该生物躯干的布局，铰链积木躯干部分如何移动，以及其模拟大脑的初始状态。西姆斯接着进行了一系列实验。他以300个具有完全随机的计算机DNA的积木生物作为实验的开始，然后，他把这些积木生物松散地放在模拟游泳池里。大多数的积木生物都是随机设计的，它们会拍打、翻滚或下沉。然而，由于偶然的机会，一些积木生物拥有了一些潜力——一些推动他们前进的动作，或一些能够驾驭自己的能力。后来，西姆斯对他的第一代随机积木生物应用了一个简单的进化准则：表现优秀的游泳者被保留，表现差劲的游泳者被去除。接下来，表现优秀的游泳生物之间发生了“电脑性爱”，相互交换部分计算机DNA，以创造出继承父母双方特征的后代生物。此外，一些后代生物还发生了随机突变——DNA发生改变。

积木生物群体中表现出了变异，任何时间点上都存在一些拥有特殊游泳能力的积木生物。模拟过程中选择最适合的积木生物并复制，从而传播推广他们的设计。这个简单的变异、选择和复制准则在大约100代的时间里不断重复。通常情况下，在20代到30代之后，笨拙、随意、摇摆、会翻滚的积木生物就开始进化成能够游泳的生物（见图9-2）。有些积木生物长出了巨大的核心躯干，背部有一条摆动的尾巴；一些积木生物的尾巴像海豚一样上下摆动，而另一些则像鲨鱼一样左右摆动。好几种积木生物长出了像鱼一样的鳍；另一些则长着又长又细的体节，像蛇一样嗖嗖作响；还有一些积木生物长出了许多小胳膊，它们可以像千足虫一样旋转，还有一些则进化成了非常优雅的海马形状。

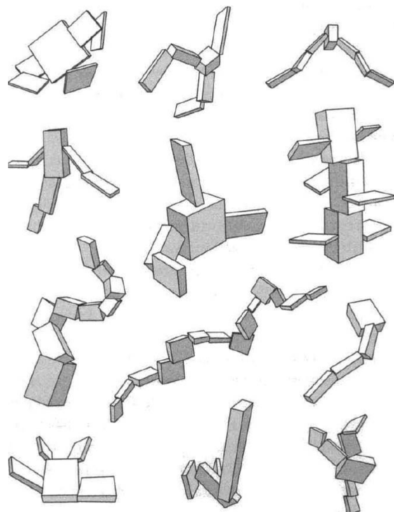


图9-2　进化后的积木生物

为游泳而进化的生物例子。

资料来源：西姆斯（1994年）。

进化算法并没有找到最佳的游泳方式，但用丹尼尔·丹尼特的话来说，进化准则发现了各种各样的“生存把戏”。[5]水的基本物理特性允许生物可以采用大量但又不是无限多的运动方式。在水中游动还需要一些能够用诸如鳍状物或流体动力学体形来稳定自己的能力，因此，水的物理特性限制了可以实现成功游动的生物的外观。这就是为什么所有真实的水生生物都包含某些目的性的变化，就像人造水上机器一样，这也是为什么西姆斯的积木生物会在进化过程中迅速“发现”这些成功的设计。

西姆斯的计算机进化模拟成功地在一个复杂又有约束条件的世界中找到了好的技巧，这还不仅仅局限于游泳问题。西姆斯还进行了一组类似的实验，在实验中，他给这些生物设定了在一个模拟重力的平坦表面上行走的目标。再一次，在经过二三十轮的进化后，随意摇摆的积木生物进化成了可以快速移动的蟹类生物、可以滑行的蛇类生物，以及能够爬行、跳跃、翻滚，甚至用腿行走的生物。在另一组实验中，为了争夺模拟食物而相互竞争的积木生物甚至进化出了胳膊、爪子和嘴巴。

西姆斯事先没有在他的程序中置入任何解决方案。虽然他创造了进化起作用的条件（例如计算机DNA和选择过程），但并没有“设计”任何关于生物自身的东西——程序中没有关于鳍、尾巴、腿或爪子的东西。所有这些设计都是在历经数代的进化过程中被发现的、自下而上产生的。[6]在数百万年的时间里，这些过程给生物带来了令人难以置信的创新

设计，比如眼球、装甲电镀、有毒喷雾、翅膀和对生拇指。

然而，进化究竟是如何创造出如此多新颖的生存把戏的呢？

一种创新算法

我多次提到进化是一种算法，但这究竟意味着什么呢？你可以将算法视为采用一些输入的食谱，例如面粉、鸡蛋、糖和黄油。通过某些过程对其进行机械加工，例如将其搅拌均匀，在175℃下烘烤15分钟，它将产生一些可靠输出，例如饼干。丹尼特以网球锦标赛的四分之一决赛、半决赛和决赛为例进行了说明——某个人或某个组织输入球员，比赛过程根据一组规则进行，然后可靠地输出结果。锦标赛赛制是一种相当通用的算法，并不局限于网球。它可以用于高尔夫球、足球、电脑游戏、弹塑料片游戏或其他任何的基质。基质可以看作算法产生作用的材料或信息。

有些算法是基质中立的，也就是说，只要有一些基本的设置条件，它们就可以被剥离成一个可以在各种基质上运作的基本核心。例如，你可以想象一个从大到小排序的算法，这样的算法可能对排序苹果桶的重量或列表中的姓氏长度有用。定义算法的不是特定的基质，而是算法用于处理信息的逻辑。排序算法实际上并不会对苹果桶或名字进行排序，相反，这个算法是根据苹果桶的重量或者名字的拼写来计算的。算法是处理信息的准则，本质上是计算机程序。[7]

进化是一种基质中立的算法，它需要获取关于事物设计的信息，并通过一个过程盲目地研究这些信息。进化也是递归的：它在一个循环中的输出会成为下一轮的输入。这种循环性意味着它会无限地运行，直到有什么东西阻止它。丹尼特举的网球锦标赛的例子也是递归的，如四分之一决赛的输出是半决赛的输入，这个过程会一直持续到确定获胜者为止。然而，生物进化并没有预先设定的停止点，它会不断地重复运转，直到太阳爆炸或者地球变得无法居住。

乐高积木图书馆[8]

在一次圣诞节的聚会上，斯图尔特·考夫曼在比赛中赢得了一套乐高积木。[9]这对考夫曼来说是一份特别合适的礼物，因为他的研究课题之一

就是观察进化是如何从简单的积木块构建出极其复杂的设计的。每个孩子都知道，乐高的巨大吸引力在于，它简单而多彩的塑料积木可以通过无数种方式拼装在一起，从而产生有趣而复杂的结构。

即使是最简单的积木块也可以附加大量的排列，例如，两个小矩形块上包括一个凹坑宽和两个凹坑长（我们将其称为1×2块），可以在14个不同的连接位置上排列在一起（假设两个积木块的颜色相同），而两个2×2块能够以33种方式连接在一起。随着积木块的数量、大小和附加块选项的数量增长，可能的排列数量会激增。即使是中等大小的乐高积木套装，也能构建出各种可能的结构，其总量之大令人难以置信。

尽管如此，乐高积木块可能的排列方式还是有限的。进化理论家把这样一组可能的排列称为设计空间。假如有500块不同形状和颜色的积木套装，它被恰当地命名为“创造者”，我们可以说，这一套乐高套装可以创造出约10^{120}种独特的乐高积木块设计。借用一些丹尼尔·丹尼特的措辞，我们称之为“所有可能的乐高设计的图书馆”。[10]我们可以想象这个比宇宙本身还要大得多的图书馆（宇宙中只有大约10^{80}个原子）里面有一排又一排的书架，上面堆满了纸质记录卡，每一张上面都有一个独特的乐高设计说明。如果我们在乐高图书馆中闲逛，随意翻看记录卡，就会发现绝大多数的设计都很无聊。会有33个设计只是由一个蓝色2×2块以不同的方式连接到一个红色2×2块，另一组33个设计是关于一个蓝色的2×2块连接到一个黄色的2×2块，数以百万计的其他方式被用于连接两个积木块，数以万亿计的方式被用于连接3个积木块，等等。

除此之外，图书馆深处的某个地方还埋藏着一件神奇的由384块积木构成的乐高宇宙飞船，它将会让任何一个7岁的孩子心跳加速。同样，这里还有一个迷人的由220块积木构成的乐高积木马，以及由405块积木构成的乐高城堡。图书馆中还有一个由406块积木构成的城堡，它与由405块积木构成的乐高城堡的设计相同，只是增加了一个积木块。事实上，所有的设计都有一个变体块。不过，尽管城堡设计有数万亿种变体，但是如此“有趣”的设计在乐高积木图书馆中极其罕见。图书馆中还有很多无聊的、随意的、莫名其妙的设计。也有很多设计根本行不通，在面对现实世界的重力时，它们会被自己的重量打翻、折断或压塌。套用丹尼特的话说，你在全世界的海洋中找到一滴水的机会，要比在乐高图书馆的所有垃圾中找到一个特定的有趣设计的机会大得多。

假设我们被赋予了一项费力不讨好的任务：在乐高图书馆中搜索有趣的设计，那么该怎么做呢？要如何在比宇宙还要大得多的乐高图书馆

中找到城堡、宇宙飞船和马呢？如果我们随便在图书馆里闲逛，偷看书架上的设计记录卡，也许要经过几百万年我们才会发现一种勉强称得上有趣的设计。我们需要一种能够在巨大的设计空间中可靠、快速地找到好的设计的算法。进化就是这样一种算法，事实上，进化本身就是伟大的胜利者。

进化设置

所有的算法都需要一些设置来完成它们的工作。因为算法是处理信息的，所以我们首先需要一些方法来将潜在的乐高设计转化为“信息”。我们可以通过多种方式来对乐高设计进行编码，比如用句子进行描述：“一个2×2的蓝色积木上附加着一个2×6的红色积木。”我们可以用建筑师创建草图的方式绘制设计图，或者也可以创建一个描述乐高设计的特殊代码，例如RED26TOPBLUE22AT56TO12，又或者在计算机上将乐高设计表示为一串1和0，例如101011100100101。具体如何编码并不重要，重要的是，它可以通过可靠的编码和解码来准确地描述这些设计，这样的设计编码被称为模式。[11]一旦我们建立了模式，就可以用这个模式来表示设计空间中的任何可能设计。

我们需要一些设备来存储模式所代表的信息。在乐高图书馆的例子中，我们可以假设其设计是写在纸质记录卡上来存储的。例如，一张卡可能记录的是RED26TOPBLUE22AT56T012，另一张卡可能记录的是YELLOW26 TOPBLUE22AT56T012。

接下来，我们需要一种机制，将模式所代表的理论设计转换为现实世界中具体的塑料乐高建造物，这时候就需要一个模式解读器。在生物世界中，模式解读器就是一种将DNA转化为生物的机制。例如，在鸟类中，受精卵能够将小鸡的DNA设计转换成一只活生生的小鸡。对于人类和其他哺乳动物来说，模式解读器就是一枚植入雌性子宫的受精卵。正如丹尼特所指出的，这也是小说和电影《侏罗纪公园》情节的致命缺陷。在《侏罗纪公园》中，科学家们用恐龙的DNA再造恐龙——他们真正需要的是一个恐龙模式解读器，换句话说，就是雌性恐龙和恐龙蛋。

生物系统的一个重要特征是模式对自身的模式解读器的设计进行编码。在人类中，当女性胎儿只有20周大的时候，她就已经发育出了卵巢和上百万个卵子。因此，在一个母亲生产之前，在她的子宫里不仅有自己的女儿，而且在她的女儿身体里面也有能生育未来的孙子、孙女们的

卵子。

然而，乐高玩具还没有发展到可以自我复制的地步。因此，假设乐高模式解读器是一个7岁的孩子，这个孩子知道模式代码，我们在孩子面前倒出一大盒乐高积木，然后递给他一张卡片，上面写着设计的代码。这个孩子会尽职尽责地捡起必要的塑料部件，将它们组装成纸条上的代码所描述的设计，我们称这个孩子为“解读者”。

我们还需要一个词来描述解读者在构建什么，需要一个术语来区分设计空间中理论的、潜在的设计和实际已经完成的设计。我们将使用进化哲学家戴维·赫尔（David Hull）提出的一个术语：互动者[12]。互动者是指从可能的设计空间中呈现出来的，并且在一个环境中实现了的“真实”设计（尽管“真实”也可以指存在于计算机上）。在一个进化系统中，一旦设计被呈现并“真实”存在，它在某种程度上就会与所处的环境产生相互作用，然后承担该环境的选择压力。

进化设置的一个必要部分是适合度函数。到目前为止，我对在乐高图书馆中寻找什么是相当模糊的。我只是说我们在寻找“有趣的设计”，但谁会对什么感兴趣呢？在卡尔·西姆斯的模型中，适合度函数是由游泳速度定义的。如上所述，设计展示了目的，由于乐高玩具的生产目的是娱乐儿童，我将假设有第二个7岁的孩子，他将担任乐高设计适合度的仲裁者。解读者将按照各种模式指定的方式组装乐高玩具，并将其交给仲裁者。仲裁者会检验各种玩具，并在0（“讨厌、无聊”）到100（“酷毙了”）的范围内打分。

进化过程：孩子的游戏

有了进化完成其工作所需的所有信息处理机制之后，我们还需要初始化系统。假设乐高套装的各个部分散落在担任解读者的孩子面前的地板上，我们从乐高图书馆随机抽取100张上面有乐高设计编码的卡片交给这个孩子。这个孩子会尽职尽责地把卡片上指定的每个玩具都组装起来，放在担任乐高设计适合度仲裁者的孩子面前。仲裁者会对每个玩具进行评分，由于这些玩具完全是随机构造的，大多数玩具的评级都为零或接近零，但也有些玩具比其他玩具有趣一点，评级略高一些。我们把仲裁者使用适合度函数标准来给玩具打分的过程称为选择。

解读者之后会拿出排名在前两位的玩具进行改造：将这两个玩具的模式卡随机切成两半并交换这两部分。改造后的新玩具具有了两款获得

高评分的玩具的设计特色。在进化术语中，这种交换被称为杂交。进化算法的一个重要特点是设计越合适，其平均变量就越多，这有助于使设计适合的特征在群体中得到放大。然而，对乐高玩具的数量也是有限制的，因为解读者可以用来构建乐高玩具的积木块是有限的。为了扩大群体中合适的玩具的特性并减少不太合适的玩具的特性，我们将实施以下规则：前20个合适的设计将通过杂交来使每对模式卡拥有4个变体，接下来的20个设计的每对模式卡将有3个变体，再接下来的20个设计的每对模式卡将有2个变体，接下来的20个设计的每对模式卡只有1个变体，而最不合适的20个设计将没有变体。我们将通过抛硬币来打破适合度的束缚。因此，我们将从原来的100个设计中制作出100个变体（40 ＋ 30 ＋ 20 ＋ 10 ＋ 0）。一旦制作好变体，我们将拆掉原来的100个设计，并将零件放回箱子里。如果一个高评级的设计需要比盒子里保有得更多的部件，我们会将最低评级的设计拆分以提供必要的部件——即使在乐高世界里也存在对有限资源的竞争。

此外，我们还假设解读者这个孩子并不完美。有时候，积木搭建过程中会出现错误。例如，他可能会错误地将模式从一张卡转录到另一张卡，并意外地更改、删除或添加之前不存在的符号。这些错误大多数时候都不会被注意到，可能只是改变了某一个部件的颜色，或者稍微改变了一个特定积木块的方向。但是，这样一个随机的错误可能会对后续玩具的适合度产生重大影响，这种随机误差就是突变。

现在，让我们想象一下这两个孩子（解读者和仲裁者）经历几十次循环过程后会发生什么。就像卡尔·西姆斯的电脑模拟一样，最初我们会看到一群随机的低适合度的玩具。解读者可能会在相当长的一段时间里费力地完成许多无趣的设计。但总有一两个设计会引起仲裁者的注意，于是，这些设计就会衍生出不同的版本。我们很快就会看到一些更适合、更有趣（至少在仲裁者眼中是这样）的设计。玩具群体的适合度水平将开始上升。我们也看到在玩具中会出现共同的特点，它们都体现出对仲裁者的口味的迎合。例如，如果仲裁者喜欢有人和动物形象的玩具，我们就会看到有四肢和脸的玩具。如果仲裁者喜欢黄色，我们会看到很多黄色的玩具。在某种程度上，我们会开始看到有复杂结构出现，也就是任何能吸引决定适合度的仲裁者的东西，比如乐高人、乐高马、乐高狗。相对而言，与在乐高图书馆中随机搜索相比，将会很快出现符合这种适合度设计的发现。

在互动者群体中出现的共同特征和花招将把我们引向另一个重要的

观点。复杂设计本质上是模块化的。[13]人类身体中有一系列令人眼花缭乱的系统、子系统和其他组件，从心血管系统到心脏，再到红细胞，而红细胞又有自己的系统和子系统。复杂的人工设计具有相同的特点，例如汽车的刹车系统、刹车本身和一个单独的刹车片。复杂设计可以被看作模块和子模块的分层集合。在进化系统中，这些系统、子系统和组件在模式中都有对应的代码片段。因此，进化构建的模式中充满了这样的模块，这些模块组合成更高级别的模块，而更高级别的模块会组合成更更高级别的模块。在生物学中，构建人类DNA图谱的是单个基因，这些基因编码了从眼睛颜色到将有毒氨转化为危险性较低的尿素的化学循环等各种事物。在乐高的案例中，我们将看到编码肢体和颜色等事物的模块。如果某个模块有助于提高互动者的适合度，那么随着时间的推移，我们将看到该模块在玩具中变得更加普遍。如果仲裁者真的喜欢黄色，我们就会看到黄色的模块出现在许多模式图中。

事实上，正如道金斯指出的那样，选择实际上并不对互动者本身起作用，它作用的是模块。[14]我们可以把每个乐高玩具看作由积木编码的特性集合。通过在仲裁者面前展示不同的组合，进化过程会测试哪些特征会吸引裁判的眼球，然后用这些特征构建更多的东西。随着时间的推移，这些特性以及为它们编写代码的模块在玩具群体中变得越来越常见。因此，进化过程并不是选择单个玩具，而是在表达“仲裁者喜欢有四肢的、黄色的东西”。

如果将第一个仲裁者撤换为仲裁者2，突然之间，适合和不适合的东西就会发生变化。之前符合前任仲裁者口味的玩具现在被仲裁者2判定为“恶心”，我们很快就会看到适合度的崩溃。但在这次崩溃之后的某个时刻，进化算法将再次站稳脚跟，并开始提出一些更好的设计。一旦这种情况发生，它将开始产生越来越多更好的设计变体，最终乐高玩具的适合度将再次攀升。如果仲裁者2喜欢飞机、汽车和绿色，我们很快就会看到玩具设计中出现翅膀、轮子和许多绿色的元素。[15]

进化是一个在巨大的可能性空间中筛选的过程。它会尝试一系列设计，然后增加可行的设计并减少那些不可行的设计，如此循环往复。没有远见，没有计划，没有理性，也没有有意识的设计，有的只是针对算法进行的盲目又机械的反复折磨。

复制因子只想复制

“但是坚持住！”你可能会哭，“这真的是没有设计师的设计吗？”算法本身可能会在没有任何帮助的情况下缓慢运行，但是所有这些设置呢？卡尔·西姆斯做所有编程怎么样呢？在乐高思想实验中安插人类解读者和人类裁判怎么样呢？

要理解内生进化的逻辑，或者说是没有外部设计者或程序员帮助的进化，最好的方法是研究一个已知的例子：地球生命的发展史。[16]地球大约有46亿年的历史，在最初的10亿年里，地球上没有任何生命，只有丰富的化学物质在冷却的海洋中徘徊，在形成的陆地中被搅动，并被火山喷射到大气中，大量分子在加热、冷却、晃动和放电时相互作用和反应着。随着所有这些相互作用和反应，一些更为复杂的分子开始出现。

最终，在大约30亿到35亿年前的某个时间点，发生了一场引人注目的意外——能够复制自身的分子被创造出来。第一个可以自我复制的分子无疑是相当简单的。一个自我复制的分子可能会从周围的化学物质海洋中吸引一些“相反的”分子，这些分子反过来又会吸引与它们“相反”的分子，从而制作出一份分子的原件。即使是一个简单的自我复制分子也不太可能是随机产生的。但是，万万亿分之一的机会也意味着是有可能的，有可能在几十亿年的时间里，有几万万亿个分子发生了几百万万亿次的相互作用。一些科学家还认为，热力学定律可能大大提高了这种可能性。早期地球上有来自火山和其他地热活动的巨大自由能量，反应网络和复杂的有机分子是新陈代谢和自我复制的基础，它们还是所有能量通过地球海洋化学渗透的可能方式。[17]

科学家还不知道第一个自我复制分子（或者说分子们）的确切组成，这是一个重要的研究领域。然而，我们确实知道，至少有一种分子看起来像是RNA的前体（地球上所有生命在核糖体的RNA上都有共同的基因——DNA是最近的进化发明）。但是，一旦第一批可以自我复制的分子从早期地球的随机化学晃动中产生了，一种奇特的逻辑就占据了主导地位。正如理查德·道金斯在《自私的基因》一书中指出的，没有复制过程是完美的。复制中出现一个错误，这里或那里发生了一些改变，就会导致自我复制的分子产生一些与自身存在细微差别的东西，但这个错误又没有大到足以损害自我复制的过程。一段时间以后，化学汤中会充满各种略有不同的自我复制分子。汤中的各种可以自我复制的分子的数量是否相同？正如道金斯所言，肯定不同。有些分子在化学上比其他分子更稳定，因此会“活”得更久，或者复制的速度更快。如此一来，寿命更长、质量更好、复制速度更快的分子将在分子群中更为普遍。此外，

在这些可以自我复制的分子所处的邻近区域里，复制所需的原料供应是有限的。如果随机突变使得一种分子在化学的应用原料方面比另一种更好，那么这种分子就能够拥有更多的自身拷贝。争夺有限资源从一开始就是进化的一个主题。

最终，一些分子可能会偶然发现一些化学结构，它们会让其他自我复制因子的结构不稳定和破裂并破坏其自我复制过程，然后使具有攻击性的分子能够将被攻击的分子释放出来的分子结合起来供自己使用。这种偶然的创新会立即带来巨大的优势，因为从同类的自我复制者那里获得预先组装好的原材料要比自己找寻并组装它们容易得多。易受干扰和化学物质被盗用的分子数量将迅速下降，而厉害的、发动干扰的分子将迅速复制自己。但一些复制分子也可能碰巧含有一些化学物质外壳，例如脂质，这些化学物质在它们的自我复制机制和外部世界之间提供了一个缓冲地带。有了对干扰物的防御，受保护的分子将更频繁地进行自我复制，并在分子群中变得更普遍。一旦达到了能够复制出对外界产生屏蔽作用的干扰分子的阶段，我们就拥有了现代人眼所能识别的某种东西——病毒。病毒可能会或可能不会被视为活着，但它肯定正在形成。

这些分子战争背后的逻辑非常简单：好的复制因子被复制。[18]这听起来像是赘述，但这正是进化最微妙和强大的驱动力之一。如果随机突变导致了复制能力的差异，那么任何促进复制的差异都会随着时间的推移在群体中变得更加普遍。进化最终会选择支持复制的模块。这也是道金斯著名的“自私的基因”理论的精髓。[19]道金斯所说的自私，并不是指人类或其他生物的基因使我们在寻求生存的过程中天生自私。事实上，合作对包括人类在内的许多物种来说是一种至关重要的生存技巧。“自私的基因”所指的是复制的逻辑，即擅长支持自己的复制的基因将被复制。任何其他策略都无法在竞争的世界存续下来。

“好的复制因子被复制”的逻辑有时会在困扰我们的电脑病毒中发挥作用。一种“成功的”计算机病毒（从黑客或病毒的角度来看）是一种可以大量复制其自身的病毒。大多数病毒会通过使用收件人的电子邮件地址簿复制。因此，病毒要被复制，就必须欺骗接收者打开电子邮件。病毒在电子邮件中出现的主题经常是“我爱你”“裸照”。我最近收到的一封特别棘手的邮件来自“管理员”，标题是“您的电子邮件账户将被关闭”。所有这些都是黑客社区为帮助他们的创造物复制而开发的花招。由于好的复制因子会被复制，如果对全世界的计算机病毒进行调查，我们将会看到这些有效的花招在计算机病毒传播中不成比例地出现。

最终，在大约38亿年前，自我复制分子的战争发展到一个点，好的复制分子拥有了看起来像现代RNA的自我复制过程，以及它们与外界之间的屏蔽膜，但是没有在高等生物的细胞中发现的线粒体、叶绿体、细胞核或其他细胞器。更确切地说，这些最初的生命形式并不比麻袋中的自我复制分子多，就像简单的蓝绿藻漂浮在海水的浮沫上。然而，一旦这些自我复制分子达到这一阶段，进化就开始了，先是出现了种类繁多的单细胞生物，这些单细胞生物又发展出了更为复杂的内部组织，可以做诸如移动、光合作用和吃掉彼此之类的事情。最终，在大约18亿年前，进化发现了多细胞生命形式的设计，导致了多样性的爆发，创造了植物、动物以及人类。

这个故事中还有许多重要的细节有待商榷，但它的逻辑中并不需要任何外生因素。只要有了模式、模式阅读器和适合度函数，我们就可以建立进化算法的逻辑。在生命的故事中，模式和模式阅读器是相同的，因为自我复制分子既是代码又是阅读器。甚至复制技术也在随着时间的推移而不断发展，最终发展出了DNA、保护DNA的核膜和有性繁殖，当然，还有无数吸引异性的策略。

在乐高实验中，适合度函数值是由外在的仲裁者决定的。在生物学中，适合度函数是内源性的。有一些决定适合度的约束因素是固定不变的，例如物理和化学定律。但是适合度的其他重要方面与系统是共同进化的。生物环境的关键部分是其他的生物，每个生物体的进化运动和反运动都会影响其他生物体的适合度。掠食性生物可能进化出更快的奔跑速度（良好的奔跑基因会增加复制的概率），但随后它的猎物可能会发展出伪装，掠食性生物的视力可能因此得到改善，它们的军备竞赛在一场无止境的共同进化竞赛中循环往复。但是随着每一个动作的产生，适合度函数值都会发生变化。作为适合度的另一个组成部分——地球的气候也在随着生命的进化而改变。例如，植物生命的发展增加了大气中的氧气含量，为需要氧气的生命铺平了道路。当然，人类的进化正在进一步改变地球上所有生命的生存环境和适合度函数值。

综上所述，进化的过程可以归结为信息处理。为了让进化站稳脚跟，进化算法需要一种信息处理媒介：用于存储、修改和复制的物质。进化在生物世界中是无法自然开始的，直到热力学和随机机会的结合创造了第一个可以自我复制的分子来存储、修改和复制分子设计。在第三部分中，我们将看到使进化在经济中站稳脚跟的信息处理媒介是口语，后来是写作。一旦建立了信息处理媒介，就可以开始分化、选择和复制

的过程。好的复制因子可以进行复制，什么是适合的既取决于环境，也取决于与其他复制者的竞争。然后，进化开始在设计空间中行进，寻找更好的复制者。

我们已经看到了进化是如何工作的，下一个大问题是，它为何如此有效？

适合度景观的探索

想象一下所有可能用DNA编码的生物的设计空间。DNA由四种碱基构成，分别为C、G、A、T，代表胞嘧啶、鸟嘌呤、腺嘌呤和胸腺嘧啶，这些碱基成对依次排列在DNA双螺旋结构中。人类基因组中大约有60亿个DNA碱基对。为了研究DNA链比人类更长的生物的进化，我们可以想象一个设计空间，它可以容纳100亿个碱基对上所有可能的DNA排列。如果我们把每个排列都写在一本厚厚的书里，把所有的书都放在一张桌子上，桌子上的某个地方有一本3 000页的书，里面有你自己的全部DNA序列，还有10^{900}本不同的书，里面有各种大肠杆菌序列，还有$10^{12\,000}$本典型植物物种序列的书等。[20]丹尼尔·丹尼特将这个囊括所有可能生物的DNA的设计空间称为孟德尔图书馆[14]。[21]就像乐高图书馆一样，孟德尔图书馆里的大部分内容都是垃圾，如果它曾经建成的话，也必定会毁于突变。虽然有大量的设计可以成功地在地球环境中生存，但相对于难以想象的设计空间的大小，它们是非常罕见的。

就像乐高图书馆一样，孟德尔图书馆是有限的，但它只在某个时间点上有限。[22]随着时间的推移，地球上生命的历史已经见证了DNA串长度的扩展，从简单的细菌到复杂的哺乳动物，因此DNA设计空间的大小也在扩展。你可以把设计空间想象成气球的表面——它在某一时刻是有限的，但它可以扩展，就像气球会在充气时膨胀一样。虽然像孟德尔图书馆这样的设计空间的有限性具有一些重要的数学含义，但对于所有实际用途来说，它也可能是无限的。在宇宙的生命周期中，人类或任何其他生物的所有可能设计都不可能得到充分的探索。

这里我想借用丹尼特的另一张图片，让我们想象孟德尔图书馆里的每一本DNA书上都有一根金属杆，[23]每根杆子的高度代表下面特定DNA序列的适合度——杆子越高，代表设计越适合。当然，所谓的适合，主要是指在特定的时间点适应特定的环境。我们将进一步假设，孟德尔图书馆中的DNA书籍是按照某种顺序排列的，它们之间只有一个“位”的距

离。例如，序列A-G-C-C-T紧挨着序列C-G-C-C-T，序列C-G-C-C-T紧挨着序列G-G-C-C-T，以此类推。因此，就在你的DNA书旁边的是你的360亿个变体，它们只差一个基因字母。[24]我们可以“一次依据一个字母”在这片地形上移动，从一本书到另一本书，通过每一个DNA序列杆的高度，我们可以看到相应的适应度变化。

在建造物测量中，我们有了一个拥有不同高度杆的巨大山地景观（见图9-3）。这种描述方式有助于我们比较哪个设计比其他的更好，生物学家称之为适合度景观，这一概念最初由进化理论学家休厄尔·赖特（Sewall Wright）在1931年提出。[25]适合度景观以可视化的方式向我们展示了设计空间中优秀设计的位置。我们可以将良好的设计视为高适合度峰值，而在设计空间的近无限处寻找优秀设计的问题可以被重新认识为在适合度景观中找到高峰。

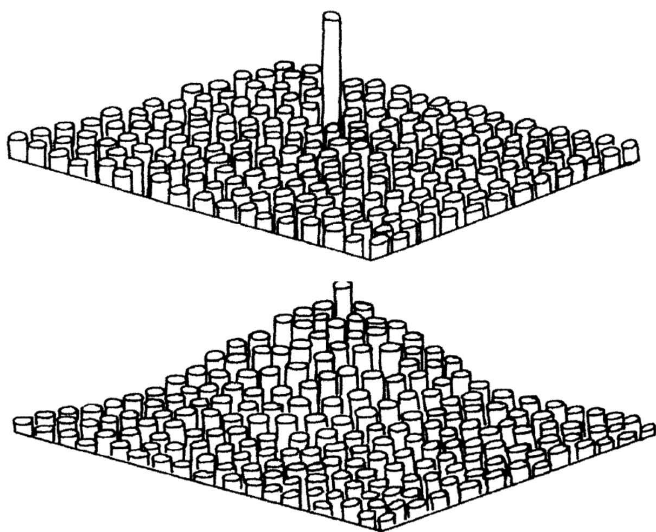


图9-3　适合度景观示意图

资料来源：丹尼特（1995年）。

这个巨大的、多山的适合度景观会是什么样子呢？斯图尔特·考夫曼指出，人们可以想象两个极端。首先，人们可能会认为每根杆子都有不同的高度，因此这个景观看起来就是完全随机的、杂乱的（见图9-4）。在另一个极端，我们可以想象一个完美有序的景观，其中的地形起伏绵延，从低适合度到高适合度，最终形成一个富士山一样的山峰高耸在景观之上（见图9-5）。考夫曼声称，实际情况介于两者之间（见图9-6）。[26]一本只有一个字母和你的DNA书不同的书，很可能和你拥有一

样的适合度，除非它是一个特别幸运或不幸的突变体。两个字母不同的DNA书还是可能和你一样具有相同的适合度。随着我们越来越远离你的DNA书，从而产生有着数十、数百或数千个字母不同的书籍，这些变化对适合度的影响会表现得更为明显。离你的DNA书越远，相应的书对应的适合度可能就越不同。随着距离的增加，变化幅度会产生较大波动，因为一些小的变化可能会产生非常大的影响，而一些较大的变化可能只会产生很小的影响。因此，任意两个相邻点之间的适合度仅仅是粗略相关的。当我们在书与书之间移动的距离越来越远时，整个景观就变得崎岖不平了。考夫曼注意到，这种粗糙的相关性在适合度景观中会创造一个更像瑞士阿尔卑斯山的地形，而不是一种杂乱无序或像富士山一样的地形。

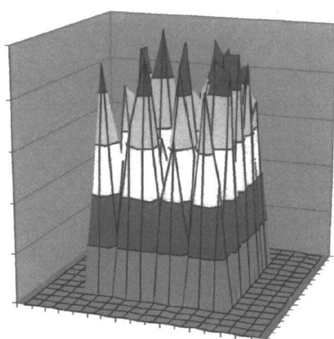


图9-4　随机的适合度景观

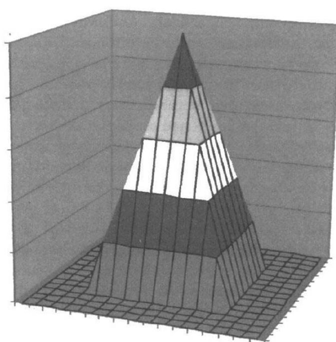


图9-5　富士山样的适合度景观

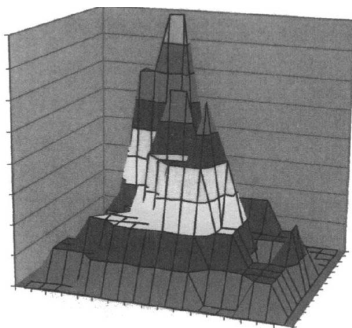


图9-6　粗相关的适合度景观

包括圣塔菲研究所的吉姆·克拉奇菲尔德（Jim　Crutchfield）在内的其他研究人员，已经研究了适合度景观的数学形状，并注意到一些其他特征。[27]与普通的山地景观不同，适合度景观中有许多平坦的斑点。正如前面所提到的，生物体DNA密码的绝大多数微小变化对适合度没有任何积极或消极的影响。因此，你的DNA书中大约有360亿个单字母变体，以及数万亿的几个字母的变体，景观周围将会有一些大而平坦的高原。这些平坦的高原内部存在裂隙，看上去就像瑞士奶酪一样布满了孔洞。进入这些洞中会导致适合度急剧下降，因为尽管绝大多数的小突变作用不大，但也有一些是相当危险的。如果从你的DNA中删除大脑中负责制作关键蛋白质的指令，那么你就会遇到大麻烦。

尽管一些小突变有很大的负面影响，但也存在带来好结果的小突变。这些小突变可以在适合度方面带来显著优势，甚至是为进一步增强适合度的突变开辟道路。克拉奇菲尔德将这些地点称为景观门户，它们是上山的快速通道，就像从山上的一个高原到达另一个高原的陆地桥梁。在前一章中，我们研究了一个模型，该模型通过生态系统中物种关系的网络结构来解释间断均衡。克拉奇菲尔德认为，平坦的高原、瑞士奶酪孔洞和景观的门户路线特征也通过在遗传变化的影响中创造非线性来促进间断均衡。大多数变化都不会产生太大影响，只有少数变化会对适合度（无论是好的或者是坏的）产生较大影响，因此这些变化可能对物种关系网络产生不相称的影响。

任何一个设计空间中大部分的小变化都不会对适合度产生大影响，只有小部分变异具有很大的影响，从而形成生物适合度景观中的相关起伏。这一点很重要，因为正是这种粗相关的特性使得进化成为搜索适合度景观中的高峰的理想算法。

大冠军搜索算法

假设我们被要求在孟德尔图书馆的适合度景观中搜寻高峰，这里需要注意的是，我们的任务并不是找到最高的山峰。整个景观中可能并不一定存在一个明确的最高峰，而是可能有多个高度相同的山峰却彼此相距甚远。更重要的是，即使有一根适合度杆比其他的杆子都略高，我们找到它的机会也非常小。而且，山峰的高度会在景观中不断变化，当我们发现最高的一根杆子时，它可能会变矮，或者另一根杆子会变高。因此，我们无法寻找某个全局最优值，只能简单地寻找在任何给定的时刻的最高峰值。

在开始徒步旅行之前，我们需要引入三个额外的挑战。第一，我们必须想象这是一个漆黑无月的夜晚。你可以感觉到你是走到了高处还是低处，但仅此而已。这代表着进化在任何意义上都无法“向前看”，它所能做的就是尝试一些东西，看看它是否有效。

第二，这个景观中毒雾附着在低适合度的山谷和缝隙中。如果你在漫游中到达的高度太低，就会掉入毒雾因窒息而死。毒雾代表了自然选择——在某些时候，如果你的适合度太低，你就会被淘汰。

第三，这个景观不是静止的，而是不断变化的。随着环境的变化，适合度函数也会发生变化，因此今天的适合度高峰可能到明天就不再是高峰了。随着环境的变化，低适合度的山谷形成新的山脉，而以前的高峰在毒雾之下塌陷为低地。在任何一个时间点上，一些地区的景观可能是稳定的，而另一些地区可能会像板块运动一样逐渐移动，还有一些地区可能是高度活跃的，就像地震带和火山爆发形成的断裂带。

一些景观的变化，比如小行星撞击地球并改变气候，代表了环境的随机变化。而大部分的起伏变化都是物种自身进化的结果。之前，我们注意到所有物种都存在于与其他物种的复杂关系网络中：掠食者、猎物、共生关系、寄生关系等。随着物种在进攻与防御、合作与竞争之间展开军备竞赛，一个物种的进化改变会引发其他物种适合度的连锁反应。

为了搜索景观，我们可能会先选择一个随机的起点，然后运用以下规则：向任意方向迈出一步；如果这一步让你向上走了，就以那里为起点再随意走一步。如果没有让你向上走，则回到之前的位置再试一次。可以想象，如果你的起始点在一个山谷里，你最初可能会在谷底随意地

移动。但你最终还是会找到一条向上的路径，爬上最近的山峰。这个规则被称为适应性步行（adaptive walk，见图9-7）。[28]虽然适应性步行在攀登单个山峰方面是有效的，但它有一个重要的局限性，即一旦你到达山顶就会停下来，被困在一个局部的最大值上。也许在离山谷不远的地方有一个更高的山峰，但是你永远也找不到，因为你必须先下山才能到达那里。这种适应性步行甚至可能使你困在场地中间的一个小山丘上。

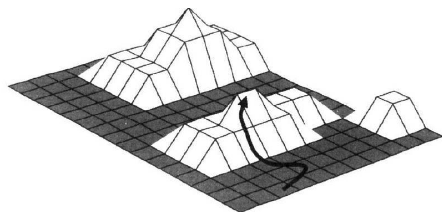


图9-7　适应性步行

另一种策略是想象你有一个非常强大的弹簧跳板。当你按下弹簧跳板上的一个按钮时，它会在一个随机的距离内向一个随机的方向把你发射出去。因此，你会连续按下按钮，希望降落在一个较高的位置，这种策略被称为随机跳跃（见图9-8）。与适应性步行相比，随机跳跃的优势在于，你可以跳过中间的山谷，从低谷跳到高峰。然而，它也有一个缺点，那就是你可能会发现自己掉进了一个死亡谷。因此，随机跳跃是一种比适应性步行风险更高的策略，因为适应性步行至少可以让你远离最低的低地并躲开毒雾。此外，随机跳跃也面临着要花很长的时间才能找到真正的高峰的问题。我们都知道，山体大致呈锥形。如果把圆锥体水平切成两半，圆锥体下半部分的表面积总是大于上半部分的表面积。因此，更多的景观表面区域都将是低适合度区域，而非高适合度区域，平均而言，随机跳跃将使你进入低适合度点，偶有幸运的例外情况。对于任何一个正在寻找适合度景观的人来说，即使是找到一个适合度高的点，总有更多的方法让你变得更糟，而不是更好。用丹尼特的话来说，在一个设计空间中，好的设计的数量总是“少得可怜”。

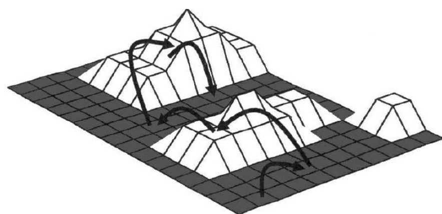


图9-8　随机跳跃

到目前为止，我们似乎面临着霍布森（Hobson）的选择。我们必须选择一种低风险，但不太可能让我们达到非常高的峰值的策略，或是一种高风险，有可能把我们带到一个高峰，但更有可能把我们带到毒雾之中的策略。

接下来，让我们尝试一下将这两种选择混合在一起的算法，这种算法采用适应性步行，以使我们在景观中爬得越来越高，同时也提供一些随机跳跃，以防止我们被困在局部的高峰上（见图9-9）。[29]我们还将权衡随机跳跃与较短距离的跳跃（跳跃距离越长，发生的概率越低），这仍然会帮助我们避免陷入局部的高峰，但降低了落入一个非常低的山谷的可能。

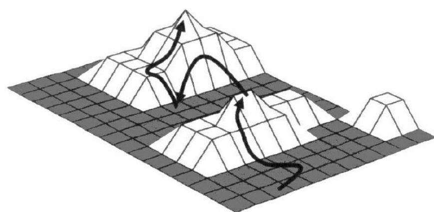


图9-9　适应性步行和随机跳跃的混合

更进一步讲，如果不是只有一个旅行者，而是有一大群旅行者来帮助我们探索景观呢？如果站在“上帝之眼”的视角，每个徒步旅行者在巨大的地形上看起来就像一个小黑点，而当一大群徒步旅行者都在寻找更高的山峰时，他们看起来就像一片乌黑的云。[30]我们会看到一片乌云缓缓爬上山坡，每个人都在进行适应性步行。最终，随着这一大群徒步旅行者展开探索，这片云将会出现在局部的高原上并随机扩散到周围。当他们继续适应性步行时，一些徒步旅行者将从“瑞士奶酪孔洞”中消失，但大多数会在局部的高原周围散布开来。我们偶尔也会看到这群人中出现了几个跳跃者，他们降落在远离这群人主体的地方。像高级侦察兵一样，这些跳跃者将在景观中寻找新的路线。一些跳跃者会从高原边缘跳跃到下面的毒雾中，一些人可能会偶然发现一条新路线，也有一些人会发现克拉奇菲尔德所说的那个门户陆桥，带他们穿过低适合度的山谷，进入一个新的高适合度的高原。

截至这里，我们就得到了一个两全其美的结果。大部分徒步旅行者采用了相对风险较低的、在景观中行进的适应性步行。也有一些徒步旅行者分散在离中心更远的地方，还有一些侦察兵拿着真正的传单，在远离人群的地方搜索。这种策略不可避免地会淘汰一些徒步旅行者，但也更有可能找到景观的高适合度区域，而不会陷入局部的高原。

这种在群体中分散赌注的行为正是进化所做的，每个生物都可以被认为是一个徒步旅行者。复制过程为适应性步行提供了动力。如果高适合度的徒步旅行者倾向于将他们的模式与其他高适合度徒步旅行者重新组合，并拥有更多的后代，那么低适合度的徒步旅行者往往会在更高的海拔地区增多，而在低海拔地区减少。但是，变异（生物学中是通过杂交和突变）确保了徒步旅行者遍布整个景观。大多数的徒步旅行者将在同一地区聚集在一起——这是一件好事，因为如果徒步旅行者还活着，他们至少知道自己是在毒雾之上。并且如前所述，下山的路总是比上山的路多，至少也会有一些离群者提供一些寻找更高的高地的新方法。

如果外围的侦察兵确实找到了一条新的上升路径，比如说通过一个门户陆桥到达一个新的高适合度区域，那么高适合度将意味着它们将更快地繁殖，在新的高适合度区域创造一片新的种群云。从上帝的视角看，我们会看到高原上嗡嗡作响的种群云——换句话说，是一个新物种。[31]

这种由变异带来的赌注分散，不仅对发现提高适合度的新方法至关重要，而且对于提高某些徒步旅行者在环境发生变化时存活下来的概率也至关重要。大多数徒步旅行者可能正生活在一个不错的高适合度的高原上，突然景观发生了急剧变化，他们所在的高原坍塌了，唯一剩下的就是那些在更远处的侦察兵了，他们必须重建队伍。同样，一些之前在低适合度地区的侦察兵可能会突然发现，随着景观的变化，他们会到达一个新的、更高的海拔。这是遗传多样性在种群中所起的重要作用：如果你不把赌注分散到整个景观范围内，当环境突然发生变化时，你就有可能失去一切。

通过创建跳跃长度的扩展，进化解决了约翰·霍兰德所说的探索与开发之间的紧张关系。[32]如果进化发现了一个高适合度的高原，大多数种群就会聚集在该地区繁殖和生长，从而利用其好运及它已经发现的成功适应。但也总会有一些徒步旅行者在更远处的黑暗中跌跌撞撞，寻找上山的新方法，以防范“过度适应”。

有趣的是，霍兰德已经证明，进化会自动在探索与开发之间取得正确的平衡。当情况良好时，也就是在找到一个高原时，进化将按比例投入更多的人口或族群资源用于开发。但是，当情况变得糟糕时，也就是当山谷中的人口或个体减少时，进化就会将更多的资源用于探索。每当进化占据了适合度景观的一个新部分，它就会押注于对未知事物的采样。就像任何下注者一样，随着进化得到更多的信息，它也想对看起来

最有希望的赌注加码。霍兰德研究出了平衡开发与探索的最佳方法，并证明了进化非常接近于达到最优均衡。[33]进化是一个赌徒，一个胜算很大的赌徒。

妙把戏、强迫移动和路径依赖

我在前文中提到过丹尼特的把戏的概念。[34]把戏是在适合度景观上的移动，这些移动并不需要经历族群灭绝的痛苦，而是非常吸引人的，它们很有可能会在景观的进化搜索中被反复又独立地发现。假设景观中有一个叫做“有眼睛的生物”的巨大高山区域，所有的DNA书中都有关于制造能探测光的传感器的指令。这种通过眼睛检测和解释光信号的能力对适合度有着重大的贡献，几乎与环境无关（除非世界变黑）。就适合度景观而言，具有“有眼睛的生物”的DNA书籍构成了一片非常大的、稳定的高适合度区域。该区域的大小、高度和稳定性使得任何进化搜索过程最终都能找到它。大而高的山区也可能有多条通向它们的进化路径，因此，把戏可能独立地出现在多个物种中，而且每个物种的设计都有所不同，例如哺乳动物的眼睛设计和昆虫的眼睛设计就是不同的。

丹尼特还提到另一个因素——受迫移动。[35]在国际象棋比赛中，棋手有时会发现自己处于一种别无选择的境地，他们被限制在每一步移动都等同于自杀的位置。同样，在一个适合度景观中，进化搜索的移动也受到特定环境的适合度的限制。也就是说，一个人可以沿着低适合度的毒雾上方的山脊跑，但如果偏离了山脊线，他就会死。一些受迫移动是由物理和化学定律的约束造成的。例如，热力学定律规定，所有生物的身体和环境之间必须有某种边界，以便在它们低熵的内部世界和高熵的外部世界之间提供一种屏障。因此，所有生物都有某种皮肤、膜、外骨骼、蛋白质外壳或其他容器。其他的强制移动可能是也可能不是永久性的，这取决于环境的当前状态或物种间的共同进化。

适合度景观造成的一个结果是路径依赖，这是我们之前在书中简要讨论过的一个概念。在进化系统中，历史很重要，你将来能去哪里取决于你过去的位置。变异可以使生物种群分散到景观的附近，但你不能随意从景观的一个部分立即转移到另一个部分（上文中讲到的随机跳跃发生的起始位置是随机的，并且短距离跳跃发生的概率较大）。假设有一个代表某种鱼类设计的山峰，鱼的生存环境正在发生变化，它的生态位正在消失，它的顶峰正陷入毒雾之中。从上帝视角来看，我们可以看到

附近有另一种目前还不存在的鱼类设计的顶峰，它高于毒雾并且看起来很稳定。然而，这两种景观之间没有可通行的道路，也没有位于毒雾之上的陆地桥梁。这些设计相距太远，以至于无法通过随机跳跃来抵达彼此。在第一种鱼类的设计与第二种鱼类的设计之间没有可持续的中间生态位，鱼类是其历史的囚徒。它的特殊道路导致了其特定高峰的局限性，其未来的选择受到了过去的限制。

数学家和进化理论家已经探索了各种不同景观的替代搜索算法。有些更适合搜索完全随机的景观，有些更适合搜索高度有序的和规则化的景观。但是对于介于两者之间的景观来说，由于它们是粗略相关的，并具有高原、洞穴和门户等复杂特征，进化是最好的算法。当景观不断发生变化，当搜索问题成为一个动态的难题，当一个人必须平衡探索与开发之间的紧张关系时，我们只能说，进化确实是伟大的冠军。[36]

剥离进化，回归基础

为了回顾我们一直在建立的基质中性的进化版本，这里列出了进化工作的必要条件：

- 有一个可能的设计空间。
- 可以将这些设计可靠地编码到模式中。
- 有某种模式阅读器可以可靠地解码模式并将其呈现给互动者，在内生进化中，模式为自身阅读器的编码。
- 互动者是由模块和模块系统组成的，这些模块是通过模式中的构建块进行编码的。
- 将互动者呈现在一个环境中。环境对互动者施加了约束（例如物理定律、气候或仲裁者），任何一个制约因素都可能随时间而改变。一个特别重要的制约因素就是互动者之间对有限资源的竞争。
- 总的来说，环境中的制约因素创造了适合度函数，其中一些互动者比其他互动者更适应。

进化的过程可以被认为是这样的一种算法，在给定环境的适合度限制下，该算法可以搜索设计空间以找到适合的设计，算法对设计空间的搜索如下：

- 随着时间的推移，模式会发生变化。模式可以被任意数量的因子改变，例如杂交和突变。
- 模式被呈现为创建族群的互动。
- 对互动者的作用是一个选择的过程，其中一些设计被适合度景观认为比其他设计更适合。不适合的互动者更有可能从群体中被移除。
- 存在一个复制的过程。适合的互动者平均具有更大的被复制概率，并且比不太适合的设计拥有更多的变体。
- 随着时间的推移，有助于提升互动者适合度的构建块会被更频繁地复制，在群体中也更常见。
- 变异、选择和复制的算法过程在种群上递归地进行，一轮的输出会成为下一轮的输入。

当算法运行在具有正确参数的适当设置的信息处理基质上时，我们可以期望看到以下结果：[37]

- 从随机性中创建秩序。从简单的随机起点出发，算法将创建出从适合度函数的角度被“排序”的、复杂“有序”的设计。所有的进化过程都是在开放的系统中进行的，因此该算法实际上是在利用能量来降低局部熵并将随机性转化为有序性。
- 发现适合的设计。该算法提供了一种快速有效的方法来搜索包含了适应的设计的巨大设计空间。在内生进化中，如果设计在环境的约束下能够生存和复制，那么这种设计就是适合的。
- 不断地适应。算法会“学习”适合度函数想要什么，并找出符合这些标准的设计。如果适合度函数发生了变化，进化就会产生反映出新的选择压力的设计。
- 知识的积累。进化过程会随着时间不断积累知识。如果我

们冻结乐高的进化过程，在特定的时间点分析所有玩具的图式卡，就可以说这些卡片中的信息反映了有关玩具历史演变的适合度环境的学习或知识。同样，DNA包含了大量的信息，这些信息与过去的生物设计有关。如果你是来自另一个星球的外星人，从来没有见过地球，却从一个地球生物身上获取了一段DNA，你就可以从这段代码中了解到很多关于地球环境的信息（假设你有一个DNA阅读器）。模式就像进化过程中的硬盘，随着时间的推移，它们会存满信息。

◎ 创新的出现。在进化过程中，算法会不断地创造出新的设计变体。从理论上讲，所有可能的设计都已经存在于设计空间中，但通过发现和呈现它们，进化将“新的”设计引入了现实世界。在乐高的例子中，进化算法毫无疑问会产生一些孩子没有预先想到的设计，并且，通过使用计算机模拟进化来设计从喷气式发动机风扇的叶片到计算机芯片的实验也催生了新颖的设计。[38]

◎ 专注于成功设计的资源会不断增长。随着成功设计在资源竞争中获胜，成功设计的数量会增长，而失败设计的数量会减少。更大的群体意味着成功的模式在物质、能量和信息方面比不成功的模式“控制着”更多的资源。然而，由于来自共同进化的网络效应与适合度景观本身的形状相结合的效果影响，增长可能不会遵循一个平稳的模式，而是遵循一个间断均衡的模式。

在具有粗相关的适合度景观中，进化能非常有效地在大型设计空间中找到适合的设计，这是因为：

◎ 进化采用的是并列搜索。实际上，每一个群体的成员都是一个单独的设计实验，所以同一时间内有很多徒步旅行者外出寻找高峰。

◎ 进化在景观上创造了一系列的跳跃。它并不只是追求那些短跳跃，因为这样可能会被困在局部最优处；它也不追求太多疯狂的长跳跃，因为这种跳跃失败的概率大于成功的概率。

◎ 进化是一个不断创新的过程。算法的递归特性从来没有停止过。考虑到景观不断变化的特性，这一点至关重要。随着进化在探索与开发之间取得平衡，可能会出现或多或少的主动搜索阶段，但是这种搜索永远不会结束。这个系统没有均衡——在进化系统中，停滞是导致种群灭绝的关键。

实际上，进化会说，“我将尝试很多事情，看看哪些可行，然后多做一些可行的，少做一些不可行的”。筛选的过程中会发生一些不寻常的事情。进化算法不仅会学习适合度函数“想要的”，还会学习在总体模式中通过学习积累的知识，因此在不断寻找越来越适合的设计时，进化过程产生了新颖性。

进化就像一场戏。演员阵容和情节都是固定的，但演员、场景和许多细节都不是固定的。进化过程可以设定在生物学、计算机模拟、乐高玩具游戏中，也可以设定在人类文化、技术和经济中。无论基质是什么，只要条件存在，所有的进化过程就会产生普遍性的结果。

从进化论到经济现实

正如前文中提到的，进化论和经济学是相互交织、相互影响的，这大概可以追溯到大约160年前达尔文的时代。尽管阿尔弗雷德·马歇尔、弗里德里希·哈耶克等许多伟大的经济学家都在努力将进化论纳入经济学，但是，他们最终还是受到两件事的限制。他们努力将对生物进化的理解映射到经济进化上，并提出了一些问题，比如，基因在经济上的对等物是什么？一组公司是一个群体吗？在经济体系中，什么构成了父母和子女？[39]通常，这些早期的努力就像瓦尔拉斯、杰文斯和其他边缘主义者一样，在隐喻推理上犯了同样的错误。在第三部分中，我们的出发点将不再是生物学，而是我们刚刚讨论过的通用算法图。现代的进化算法主张进化系统是一个具有普遍规律的通用类别，我们需要搞清楚的是经济是否属于那个类别，或者是否受这些规律的约束。如果答案是肯定的，那么我们就可以说经济世界和生物世界都是这一通用类别的成员。它们可能在算法的实现上会有很大的不同，因此询问父母和后代在经济学中的含义可能是毫无意义的。尽管如此，这两个世界仍然遵循着相同

的进化系统的一般规律，因此这也解释了强烈的家族相似性（请原谅这个比喻）。

或许，这条道路上的首批开拓者是理查德·纳尔逊和悉尼·温特，他们创作出了影响深远的《经济变革的进化理论》。但从他们在20世纪70年代末和80年代初的优势来看，基质中立的进化算法理论才刚刚形成。约翰·霍兰德具有里程碑意义的《自然和人工系统中的适应》（*Adaptation in Natural and Artificial Systems*）出版于1975年，理查德·道金斯的《自私的基因》出版于1976年，约翰·梅纳德·史密斯（John Maynard Smith）的《进化与博弈论》（*Evolution and the Theory of Games*）出版于1982年，斯图尔特·考夫曼的《秩序的起源》（*Origins of Order*）直到1993年才出版。这些作品中的观念为20世纪80年代中期以来蓬勃发展的研究项目提供了种子，这些重要的工作直至今天仍在继续进行。[40]

但是纳尔逊、温特以及他们的前辈的理论中也缺少了另一个重要的组成部分——大量廉价的计算能力。正如我们所看到的，进化是一个在大群体中被数千次、数百万次甚至数十亿次无情地反复验证的算法。这样一个过程几乎不可能采用大多数传统经济学家使用的分析风格来建模，直到20世纪90年代，计算机的出现使一切成为可能。

我们现在已经有了进化这出戏的角色和情节，并且已经准备好开始挑选戏剧角色了，还建立了组织、市场和国家经济的戏剧环境。这出戏的剧本就是创造财富的故事，用莎士比亚笔下的福斯塔夫（Falstaff）的话来说，现在是“演完这出戏”的时候了。[41]

第三部分

财富起源于进化：复杂经济学中的三种进化力量

THE ORIGIN OF WEALTH

THE ORIGIN OF WEALTH

这样的生命观无比壮阔……从如此简单的起源，就能进化出最美丽、最奇妙的事物。

——达尔文《物种起源》

10

力量一：
商业设计

两人因犯罪被捕，但警方缺乏充分的证据证明哪个人是有罪的。审问者将两个嫌疑人分别关在不同的牢房里，这样他们就无法相互交流。审问者告诉每个嫌疑人，如果他能指认自己的同伴，他就会因提供线索获得奖励并被释放，但前提是同伴没有指认他。如果嫌疑人和同伴相互指认，他将会被关进监狱，不过他会因提供线索获得减刑。如果嫌疑人保持沉默，他就面临着两种可能：一是他的同伴也保持沉默，他们都会因证据不足而被释放。二是嫌疑人保持沉默，而同伴指认了他，那么他将被监禁很长一段时间。

你可能意识到了，这就是囚徒困境博弈——一个无处不在的经济问题，复杂经济学的先驱之一、密歇根大学的罗伯特·阿克塞尔罗德曾经称之为“社会科学研究的大肠杆菌”。[1]正如生物学家通常从细菌、果蝇和其他简单的有机体入手，然后才转向复杂的人类，我们将从一个简单的模型开始，逐步建立起对经济运行方式的直觉。在本章中，我们将使用囚徒困境博弈作为大肠杆菌模型，并了解如何使用进化算法来进化简单游戏中策略的设计。第三部分结束后，我们将从游戏过渡到经济整体的进化。

囚徒困境博弈

囚徒困境博弈中的选择可以用一个给出不同决定组合结果的回报矩阵来描述（见图10-1）。结果显示，两名嫌疑人都有强烈的指证动机。如果你不知道同伴会做什么，并且没办法与他协调，最好的选择是去指证。具体来说，如果你的同伴保持沉默，那么你去指证更有利（你获得的回报是1：0），而如果你的同伴指证了，指证仍对你更有利（你得到的回报是-2：-5）。因为两名玩家都面临着相同的选择条件，游戏中会出现一个均衡的解决方案，两名嫌疑人都像小鸟唱歌那样指证对方，然后都进入监狱。从嫌疑人的角度来看，这当然不是最佳结果，因为如果

他们能以某种方式协调行动，那么他们都能获得自由。

		嫌疑人 2 保持沉默	嫌疑人 2 指证
嫌疑人 1	保持沉默	0, 0 →	−5, 1 ↓
	指证	1, −5 →	−2, −2 （指证均衡）

图10-1　囚徒困境博弈中的回报

囚徒困境博弈代表着一种典型的均衡：是追求狭隘的个人利益，还是尽量和他人合作获取更大的利益？这种权衡经常出现在核武器控制、商业战略和婚姻关系等多种情况下。囚徒困境是博弈论问题的一个例子。顾名思义，博弈论适用于研究这样的情况：一个人可以定义一组玩家（例如两个嫌疑人）、一个目标（获得奖励或避免坐牢）、一个决定（指证或不指证）以及一组能将决定反映到玩家回报中的规则。[2]

囚徒困境特别有趣，因为它是一个非零和博弈的例子。在一个非零和博弈中，两个或两个以上的人合作会给群体带来更大的收益。在囚徒困境中，两个玩家总回报的最大化要通过彼此的合作来实现。也就是说，如果两个玩家都保持沉默，他们的总回报是0，而不是另一种选择中的-4。“如果你帮我梳毛，那么我也帮你梳毛”这句话就展示了一个非零和博弈的例子，因为合作能使我们获得自己无法获得的收益。相反，在一个零和博弈中，对于每一个赢家而言，一定有一个输家，一个人的收益就是另一个人的损失。例如，两个人对足球比赛的结果打赌，一个赢了，另一个就会输，不可能两个人都是赢家。我们将看到，非零和博弈的存在以及由此产生的利己与合作之间的矛盾在复杂经济学中起着核心作用。

囚徒困境给我们带来了一个难题。经济依赖于合作活动，人们需要一起工作来生产东西，进行贸易。然而，当人们只能看到狭隘的自身利益时，他们会产生一种在工作中懈怠、在贸易中欺骗的动机。即使人们并没有真正的欺骗，在不知道别人会做什么的情况下，率先为更大的利益伸出合作之手仍有可能面临风险。在《微观经济学》（*Microeconomics*）一书中，圣塔菲研究所的经济学家塞缪尔·鲍尔斯（Samuel　Bowles）用他与印度小村庄帕兰普尔（Palanpur）的农民进行过的一次讨论说明了这个问题：

帕兰普尔的农民会在预计能达到最高产量的播种日期之后几周再播种冬季作物。农民并不怀疑早些播种会有更好的收成，但没有人愿意成为第一个播种的，因为种在空荡荡的土地上的种子会被鸟吃掉。我询问是否有一大群农民或亲戚同意早些播种，然后都在同一天播种来减小损失。“如果我们知道如何做的话，”他抬头看着我说，“我们就不会这么穷了。”[3]

这种困境被称为协调问题。当人们不合作时，博弈论者称之为背叛者。那么我们应该如何在经济学中实现协调，如何避免背叛陷阱呢？

如果嫌疑人作出了决定，看到了他们的结局，然后又有了重新做决定的机会，结果会不会改变呢？你可能会认为，两名嫌疑人在监狱里待了几次之后，最终会发现他们可以通过相互合作并保持沉默来做得更好。同样，如果帕兰普尔的农民尽量协调他们的种植，而一些农民没有守信，那么这些背叛者将在下个种植季受到惩罚或制裁。

事实上，这个逻辑是有问题的。如果游戏是重复的，并且有固定的回合数，那么两名参与者就又会回到背叛的状态。在囚徒困境中，如果两名玩家都知道，第五轮是最后一个回合，逻辑就会迫使他们把这当成一个单轮游戏去玩。第五轮之后就没有了，所以在第五轮，两名玩家都会自动背叛。同理，他们在第四轮也面临着同样的困境。既然知道无法在第五轮中获得合作的奖赏，那为什么不在第四轮也背叛呢？如此一来，第一轮也会发生同样的情况。

但在现实世界，我们通常无法准确地提前得知一场游戏会进行多少回合。要是游戏的回合数有限但未知该怎么办呢？一个帕兰普尔的农民知道他一生中的种植季有限，但又不知道会有多少季。在这种情况下，游戏就会失去均衡，传统经济学分析也没有告诉我们最佳策略是什么，我们必须要用不同的方式来思考这个游戏。[4]

冠军的锦标赛

20世纪70年代末，罗伯特·阿克塞尔罗德尝试用一种不寻常的方法来回答如何重复地玩未知次数的囚徒困境游戏。他决定进行一个实验，这个实验自那之后变得十分出名。[5]阿克塞尔罗德没有用数学分析来解决这个问题，相反，他举行了一场比赛。阿克塞尔罗德让来自世界各地的14名社会科学研究者提交一个拥有最佳策略的候选人，然后在一场循环锦标赛中让不同的参赛者相互对决。其中一些策略非常精细，采用了复杂

的数学公式。但是赢得比赛的策略非常简单。它是由多伦多大学心理学和数学教授阿纳托尔·拉波波特（Anatol Rapoport）提交的。拉波波特的战略被叫做“以牙还牙”，它的第一步是合作，从那之后，参赛者只是简单地观察对手的最后一步，然后重复对手的动作。如果对手合作，那他也合作；如果对手背叛，那他就背叛。阿克塞尔罗德对这一简单策略的成功感到惊讶，并举办了第二次更大的比赛来进一步测试它。这一次有来自经济学、数学、物理学、计算机科学和进化生物学领域的62名参赛者，“以牙还牙”又赢了。

阿克塞尔罗德十分好奇：这样一个简单的策略怎么能接连打败那些更为复杂的设计呢？“以牙还牙”真的是最好的，还是有其他更好的策略等着被发现呢？此外，虽然“以牙还牙”的策略总体上是成功的，但它在应对某些策略上的性能也存在不如其他策略的情况。从某种意义上讲，它相当脆弱。想象两个采取“以牙还牙”策略的玩家相互对抗，事情进展得非常顺利，突然一名玩家犯了一个随机错误，展现了背叛行为，他们就会陷入恶性循环，永远相互背叛。这个灾难性的结果仅仅是因为一个简单的错误，它给阿克塞尔罗德感兴趣的领域——核武器控制，带来了令人担忧的影响。

阿克塞尔罗德还想探索其他策略，但又不想举办另一场更大的锦标赛。随后，阿克塞尔罗德在密歇根大学的同事约翰·霍兰德向他介绍了20世纪70年代中期霍兰德发明的在计算机上模拟进化的技术。阿克塞尔罗德决定尝试霍兰德的模拟进化算法，看看他是否可以简单地在计算机上模拟进化策略，而不必让人类为囚徒困境博弈提建议，然后让大自然主导最适者生存的战争。

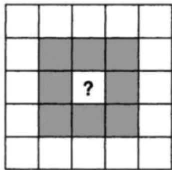
计算机模拟策略

阿克塞尔罗德于1987年发表了他的研究成果，他结合进化模拟和博弈论开展了一系列富有成果的研究。[6]克里斯蒂安·林格伦（Kristian Lindgren）建立的模型对这项研究作出了贡献，他是来自瑞典查尔姆斯理工大学和哥德堡大学的一名物理学家。[7]林格伦的模型特别有趣，他没有让成对的参与者在锦标赛中相互对抗，上演囚徒困境，相反，他让多个参与者同时进行游戏，如同糖域中相互作用的参与者那样。

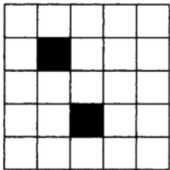
林格伦对不断演变的囚徒困境的模拟为我们提供了一个相对简单的大肠杆菌模型，该模型捕捉了进化如何在经济的复杂适应系统中展开的

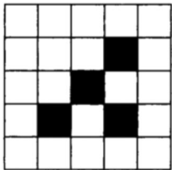
几个关键方面。第一，合作与竞争之间的内在权衡产生了一种中心矛盾，使系统持续处于不平衡状态。这种矛盾导致了合作结构的转变，这种结构会随着时间的推移自我组织、成长和解体，就像在现实经济中一样。第二，与我们在第6章的讨论一致，林格伦的模型中面对囚徒困境的参与者没有办法全局优化他们的行为。他们所能做的就是审视自己的情况和历史，然后尽力而为。第三，正如前面所讨论的，参与者的动态交互通常会导致复杂的紧急行为模式，其中许多行为模式是不可预测的，林格伦的模型也不例外。第四，也是最有趣的一点，这是一个创新的模型。就像卡尔·西姆斯探索能游泳的生物的进化一样，林格伦利用进化来寻找囚徒困境中的成功策略。实际上，我们可以把林格伦模型看作利用进化探寻所有可能的囚徒困境策略的巨大设计空间。

林格伦模型的整体结构实际上结合了两个游戏：囚徒困境和生命游戏。生命游戏是由数学家约翰·霍顿·康威（John Horton Conway）开发的，这个游戏在一个像是大棋盘的二维网格上进行（见图10-2）。网格上的每个方块都代表一个小细胞，每个细胞在游戏中的任何一点都可以打开或关闭。如果一个细胞处于打开状态，它将变为黑色；如果处于关闭状态，它将变为白色。每个细胞有8个相邻的细胞，4个紧邻的细胞和4个位于对角线上的细胞。一个细胞是打开还是关闭取决于其相邻细胞的状态，每个细胞都遵循一个简单的规则：计算有多少相邻的细胞处于打开状态。如果数字是2，那么中心的细胞将在下一轮保持它现在的状态。如果数字是3，那么中心的细胞将在下一轮中打开，不管它现在的状态如何。在其他情况下，细胞将会在下一轮关闭。



每个细胞有 8 个相邻的细胞


如果有两个相邻细胞处于打开状态，
那么细胞在下一轮会保持现有的状态


如果有 3 个相邻细胞处于打开状态，
那么细胞在下一轮会打开

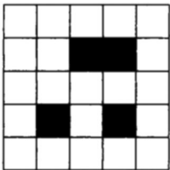

其他情况下，
细胞在下一轮会处于关闭状态

图10-2　生命游戏

生命游戏之所以得名，是因为它用简单的规则建立了细胞网格，人们能看到细胞组成的图案，有时一闪一闪就好像细胞是活的。有时网格看起来像是一团乱闪的光，有时细胞在复杂的图案中自我组织，看起来像是在培养皿中生长的细菌。你可能会注意到生命游戏和考夫曼的布尔网络之间存在相似性，两者都属于细胞自动操作，这是约翰·冯·诺依曼在20世纪60年代开发的一类高度通用的计算系统。[8]

林格伦在他的模型中创建了一个生命游戏，但是他没有按照生命游戏的正常规则让格子上的细胞闪烁，而是让它们根据细胞之间进行囚徒困境博弈的情况打开或者关闭。我们把网格上的细胞称为行为者，每个行为者都会与最近的4个邻居玩囚徒困境的游戏（见图10-3）。每个行为者都有可能采取不同的游戏策略。例如，一个可能采取“以牙还牙”策略，另一个可能正相反（总是做与对手相反的事），一个可能总是合作，等等。林格伦根据占领细胞的行为者所采用的策略给每个细胞染色，他获取了每个行为者在4场游戏中的分数，计算了它们的平均分。一轮中平均分最高的行为者被称为胜利者。然后，胜利者会像病毒一样占领邻近细胞的中心位置，然后将其获胜策略注入这个细胞，实际上是进行了自我复制。[9]例如，一个反“以牙还牙”的行为者赢得了一轮比赛，中间的玩家就会被注入反“以牙还牙”策略，然后在下一轮中使用这个策略。



图10-3　将生命游戏和囚徒困境博弈结合在一起

那么，网格上的每个行为者是如何决定使用哪个策略的呢？我们可以在网格中随机设定不同的策略，看看会发生什么。这样做会产生一个相当无聊的结果，因为最终格子会进入某种均衡状态，或者重复循环某种简单的模式。这对林格伦寻找比“以牙还牙”更好的策略的问题帮助不大，因为通过测试随机产生策略找到好东西的概率很低。

林格伦转而决定让进化搜索所有可能的囚徒困境博弈的设计空间，寻找高分策略。网格上的每个行为者都被赋予一个概要，一个由1和0组成的计算机DNA字符串，用来编码策略。林格伦认为行为者会通过归纳

作出战略决策，也就是说，他们回顾了游戏历史中的行动和反抗行动，并利用这一历史模式预测在下一轮中获得成功的策略是什么。因此，输入每个行为者的计算机DNA都是历史中的行动和反抗行动，输出的是是否合作或背叛的决定。

林格伦初始化游戏时设置的是只能记得前一行动的行为者。对于一个只记得前一行动的行为者来说，可能采取的策略只有四种：

1．永远背叛，永远背叛，不管对手做了什么。

2．永远合作，永远合作，不管对手做了什么。

3．以牙还牙，总是做对手做过的事情。

4．反以牙还牙，总是做和对手相反的事情。

在游戏开始的时候，林格伦给每个行为者随机赋予了这四种可能策略中的一种。然后，他借助随机突变周期性地破坏它们的DNA，使其经历进化变异的过程。林格伦的模型采用了三种类型的突变。第一种，点上的突变，仅仅是行为者计算机DNA中的某一单一位置发生了改变，比如01变成了00。第二种是基因复制，DNA中的一段被复制然后增加到了尾部，比如01变成011。基因复制的一个效用是行为者的记忆会扩充，从而记住更多的历史行动。因此，行为者不仅能看到前一步的行动，还可能看到前两步或三步的行动。有了更多的记忆，行为者就能想出更复杂的策略，“如果我的对手背叛了，那么我也背叛，然后我的对手会再次背叛，我就会合作”。第三种，分开的突变，DNA的一段或另一段被砍掉了，比如011101000110001变成了011011。这将缩小内存大小，从而降低策略的复杂性。

除突变外，林格伦还会让行为者犯错来制造随机的改变。一个行为者的策略可能表明要合作，事实上它却会“意外地”背叛。

一片比特雨林

为了开始这个游戏，林格伦制作了一个游戏板，板子中有16 384（128×128）个方格，每个方格里都有一个行为者，每个行为者在初始化时都会采用四种基本策略中的一种。林格伦按下开关，游戏开始，进化过程随即发生。在游戏过程中，策略会发生突变，最适合的策

略得到最高分数，并开始注入相邻的方块中。

一个生态系统诞生了。起初，随着四种策略相互作用，由随机分散的行为者构成的风景开始相互渗透——“以牙还牙”和总是合作的行为者爱上了合作，共同取得了高分。同时，总是背叛的行为者无情地利用了可怜无助的反以牙还牙者，总是合作的行为者一直合作，总是背叛的行为者持续背叛。很快，大量不合作而总是背叛的行为者集合被赶出了板子，而合作之岛出现了。这些岛屿通常会有一群总是合作的行为者处在中间位置，它们疯狂地合作，提高分数，把邪恶的总是背叛者留在海湾里。然而，这些合作的殖民地并不稳定。有时，入侵的背叛者会打破合作之岛，把每个人都拖垮。另一些时候，一个合作之岛会像细菌群一样迅速扩张，把它的边界推向充满敌意的背叛之海。最开始混乱不堪的板子，随着不同的策略在合作与背叛间的大战中重新占领不同的位置，出现了方块、之字形、螺旋形以及其他形状的图案。

很快，动荡的策略困境中又发生了其他的事情——创新出现了。基因上的突变使行为者的记忆容量扩增，从而使行为者能够回溯更多的历史，想出更为复杂的策略。很多突变体是很快就消失的无意义策略，但是总的来说，更多的记忆是一个巨大的优势，成功的新策略开始出现并复制。林格伦将一种策略命名为傻瓜策略，即如果行为者与对手在上一轮中采取了同样的行动（不管是合作还是背叛），那么行为者在下一轮就合作；如果行为者与对手采取了不同策略，那么行为者在下一轮就背叛。另一个是复仇策略，一开始双方都会选择合作，但是之后如果对手背叛了，行为者就会接连背叛，目的是在再次合作之前让对方明白它的想法。就像之前提到的，“以牙还牙”的一个问题就是如果两个玩家相互以牙还牙，其中一个不小心犯错背叛，两个行为者就会永远相互背叛下去。还有一个是公平策略，一开始双方是合作的，但如果其中一方背叛了，采取公平策略的另一方就会“生气”，然后用背叛惩罚对手，直到对手重新开始合作。然而，如果采取公平策略的行为者犯了错，它就会通过合作“道歉”，直到被对手原谅，双方再次开始合作。因此，公平策略往往在嘈杂、容易出错的环境中表现很好。

森林之王

那么谁是赢家呢？最好的策略是什么呢？林格伦发现这是一个荒谬的问题。在像林格伦模型这样的进化系统中，事实上并没有单一的赢

家，没有最优的、最佳的策略。相反，任何在特定时间点活着的玩家都是赢家，因为其他玩家都死了。要想生存下去，一个行为者必须有一个策略并为之奋斗，以某种方式谋生，抵御竞争对手，应对变幻莫测的环境。“每个人都是赢家”听起来像是逃避，但生存并不是什么了不起的壮举。正如斯图尔特·考夫曼指出的，地球上曾经生活过的绝大多数物种现在都灭绝了。在任何一个时间点上，只有适合的风景中极其微小的一部分能够成为现实。

“谁赢了”是一个无意义的问题，一片森林就是一个由植物、昆虫、鸟类和动物组成的高度复杂的生态系统。这些生物都是使用不同生存策略的行为者，有的相互合作，有的相互竞争，那么谁是森林之王呢？谁有生存繁殖的最佳策略呢？是现代人吗？人类似乎认为自己在主导这场演出，但我们只是进化历程中的一个瞬间，从长期来看，人类的未来也是不确定的。那小田鼠呢？这种生物比人类存在的时间长，数量也比我们多。在地上疾行的蟑螂呢？蟑螂存在的时间甚至比老鼠都要长，并且似乎能在一切情况下存活，从杀虫剂到核爆炸。那些我们努力消除却不受影响的杂草呢？或者那些快乐地在森林中所有哺乳动物肚子里搭便车的大肠杆菌呢？另一个更为复杂的问题是，如果你做了一项调查，研究所在地的森林，用某种方法给前100种成功的生存繁衍策略排名，然后再看看100年后的清单（在生物学的历史中仅仅是一瞬间），它看起来会完全不同。在今天成功的策略可能在100年后非常失败，百年后最成功的策略现在可能只是一般成功，甚至都不存在。

同样，我们也不能说囚徒困境生态系统中的哪种策略是赢家。林格伦的模型表明，偶尔可能会有一个特定的策略出现，在游戏中主导一段时间，在阳光下度日，但最终还是会被创新竞争者打败。有时会有几个策略共同崭露头角，争夺游戏板的“市场份额”控制权，然后一个外来者会进来把它们全部打败。在其他时候，两种策略会作为一对共生体同时出现，但如果其中一种陷入困境，两种策略就都会崩溃。游戏中也有蟑螂，它们会采取一些简单的策略，比如“以牙还牙”，这些策略从未占据主导地位，但似乎无论如何都能存活下来。

游戏也以间断均衡的模式进行着。时间久了，一个确定的策略秩序就会出现，游戏就会进入几个策略主导的相对稳定的阶段。在这些阶段中，每个参与者都单独行动，行为者的策略相对稳定，没有一方有特别的动机去改变，从而确立了约翰·梅纳德·史密斯所说的进化稳定策略。[10]然而，小的创新迟早会出现，积聚动力，突然间整个板子都会进入火热

的变革期，旧的秩序被摧毁，新的秩序建立起来。有时，这种模式的繁荣时期有利于合作战略，在这期间，每个人都能获得巨大的利润。然后，在原因不明的情况下，经济的乌云聚集起来，一个充斥着背叛者的时代将会到来（见图10-4）。

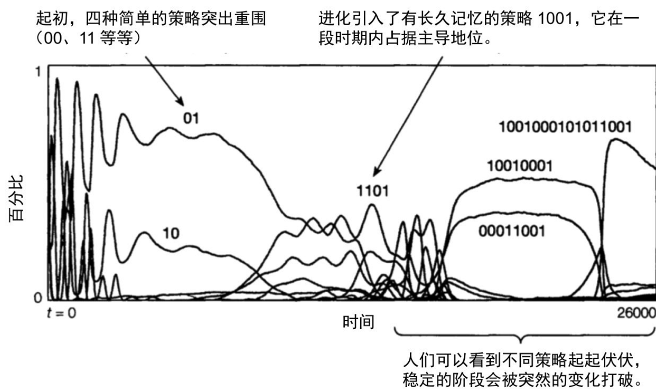


图10-4　随着时间变化的策略总数

资料来源：林格伦（1997年）。

无法预测但能够理解

所有这些动荡的出现，都是因为任何一个策略的成功或失败都高度依赖于特定时间点环境中的其他策略。人们可以把这个板子想成一个巨大的生态网络，就像第8章中贾因和克里希纳的生态系统模型，系统某部分的小变化能够激起涟漪，在别的地方产生巨大的变化。例如，如果一个总是合作策略的岛屿被“以牙还牙”策略的保护层包围着，而总是背叛的参与者只能待在外面，无法破坏总是合作策略。然后，一些小突变使“以牙还牙”策略进入了一个向下的漩涡，为总是背叛策略创造了一个侵袭入口，合作之岛很快就倒塌消失了。

这种对微小变化的敏感性意味着人们不能用方程来预测模型的结果。它和实体经济一样“不可预测”。人们可能会认为，这种无法预测性是由模型中的随机因素造成的，即行为者行动中的随机突变和随机错误。然而，作为一个计算机程序，这个模型实际上是完全确定的。如果生成用于统计突变率和错误的随机数表，然后在第二次运行中重用该表，就会得到与第一次运行相同的结果。你可以知道所有行为者确切的起始位置、行为规则，甚至能提前知道游戏中所有的“随机”数字。然而，即使有了所有这些信息，我们也没办法预测模型的精确行为，因为

这太复杂了。

要想知道发生了什么，唯一的办法就是运行这一模型使其进化，没有捷径。尽管无法预测这样一个模型的行为，但从科学的角度来说，我们可以理解它。例如，我们可以看到模型有一些特定的调整参数，其中包括变异率、错误率以及游戏中的相对收益，这些都会影响游戏的宏观行为。对于这些参数的某些值，我们可能会得到高度合作、高分的结果。其他值可能会使模型进入一个很难摆脱的沮丧、不合作的状态。我们还可以看到，与其他值相比，一些值可能会在新策略中产生更多的创新。因此，即使我们无法作出具体的预测，也可以作出“0.001或以下的突变率导致低创新和低合作”或“突变率在0.001至0.01之间，环境充满活力和创新，具有高度合作和高利润的特征”等陈述。同样，一些参数值可能比其他变量提供对某些策略更有利或不利的环境。因此，我们可以作出一些可能的陈述，比如“在有很多干扰的情况下，友好但强硬策略在高度合作的环境中表现更出色，而铁拳策略在低度合作的环境中表现更出色”。我们也可以用相同的参数数千次地运行模型，但是使用不同的随机冲击来观察结果的稳定性。然后我们可能会说：“对于x的参数设置，高利润环境出现的概率为60%，低利润环境崩溃的概率为40%。”因此，我们虽然无法预测模型的具体结果，但通过探索参数空间和收集统计数据，仍然可以很好地了解模型的行为。在许多方面，深入了解系统的工作原理最终可能比能够作出预测更有价值。

在前一章中，我们看到了卡尔·西姆斯如何利用进化来找到解决游泳问题的好方法。在林格伦的模型中，我们看到了如何利用进化来寻找经济问题中的有效策略。结果表明，没有最好的战略。相反，进化过程创造了一个策略生态系统，一个在创造性破坏中随时间变化的生态系统。我们现在可以问，如果进化可以搜寻简单经济游戏的设计空间，它也可以搜寻更复杂的游戏的设计空间吗？它能搜寻超大型游戏——36.5万亿美元的全球经济的设计空间吗？

巴别图书馆

全球经济的设计空间听起来像是一个荒谬的想法，而我们还要简单地考虑另一个荒谬的想法——所有可能的文学的设计空间。丹尼尔·丹尼

特在他的《达尔文的危险思想》（*Darwin’s Dangerous Idea*）一书中要求我们想象一个无限大的图书馆，里面包含着用英语写的所有500页的书。[11]他根据豪尔赫·路易斯·博尔赫斯（Jorge Luis Borges）的一本小说，把这个想象的图书馆称为巴别图书馆。一页可以容纳大约2 000个字符，所以一本500页的书可以容纳100万个字符。英语中大概有100个字符，包括小写字母、大写字母、数字和标点符号，这意味着巴别图书馆里大约有$100^{1\,000\,000}$本可能的书。就像孟德尔图书馆和乐高图书馆一样，巴别图书馆比宇宙还要大很多，因此只能是一个理论建构。

想象一下，我们可能在巴别图书馆中找到什么？有一本书中全部是空白页；有一本书中只有一个字母a，剩下的都是空白页；还有一本书500页里全都是问号。图书馆里的大部分书都是无意义的字符的集合。正如丹尼特所指出的，找到一本含有一个语法正确的句子的书“可能性微乎其微”。然而，这个图书馆还保存着一本完美的《白鲸》副本，以及威廉·莎士比亚的完整作品。它还包含一本除了鲸鱼被称为“波比·迪克”外，在所有方面都与《白鲸》一模一样的书。还有一本鲸鱼被称为“科比·迪克”和一本鲸鱼被称为“多比·迪克”的书。一个奇怪的地方是，图书馆里的某个地方还有一本500页的关于你的生活的完美传记，包括还没有发生在你身上的所有事情，例如你的死亡。

如果一本书不足500页，比如说阿尔贝·加缪（Albert Camus）的《西绪弗斯神话》（*The Myth of Sisyphus*）只有192页，那么将会有一本500页的书，前192页是《西绪弗斯神话》的精确副本，其余308页是空白的。同样，图书馆里超过500页的书会通过多册来呈现，例如詹姆斯·乔伊斯（James Joyce）的《尤利西斯》（*Ulysses*）有933页，可以分成第一册和第二册，第一册500页，第二册433页，后面有67页空白页。人们甚至可以说所有语言的书都在书架上的某个地方，因为图书馆里有用各种语言写作的书的英文译本。同样，尽管还没有写完，但2042年最畅销的书也放在图书馆里。

有些读者可能会接受“能够创造出特定乐高组合的乐高设计是有限的”，而回避“能写出来的书的数量是有限的”这种说法，即使它是一个穷尽几个宇宙都探索不完的超越天文数字的量级。然而，这种设计空间确实是有限的，这是一个数学事实。我们的直觉是，随着新作品的诞生，这个设计空间会随着时间的推移而扩大。设计空间的概念包含一个悖论：一方面，它们是有限的，另一方面，它们是无限的，可以像前面提到的气球表面一样扩大。唯一能使巴别图书馆的图书数量确定的是丹尼

特限制每本书都是500页，但这是一个随机的限制。理论上，一本书的页数并没有固定的限制。[12]然而，总有一本书是特定的时间点上最厚的书，从而定义那个时刻书的厚度上限。现在，根据《吉尼斯世界纪录大全》记载，有史以来字数最多的书是马塞尔·普鲁斯特（Marcel Proust）的《追忆似水年华》（*Remembrance of Things Past*），它共有13卷，总计9 609 000个字符。然而，在普鲁斯特写这本书之前，书的字符上限是比这个低的，将来很可能有人会写一本拥有超过1 000万个字符的书。但没有人会写或读一本无限长的书，因为物理定律不允许概要阅读器读取无限长的代码。[13]现实世界中呈现的任何设计都必须具有有限的概要，而有限的概要意味着有限的设计空间。

设计空间是有限的，但是可以随着最长可表达概要外层的增大或减小而扩大或缩小。正如上一章提到的，地球生物的设计空间会随着DNA不断变长以及物种表达能力的增长而扩大。同样，我们也能想象图书馆的设计空间随着时间发生改变。当人们将象形文字刻在石板上时，书籍无疑变得更短了，也许在未来，当书籍可以直接下载到我们的大脑中时，普鲁斯特的书也会显得很简洁。

巴别图书馆说明了有关设计空间的另一个要点：我们可以为能用一系列符号表示的任何东西构建一个设计空间，这些符号可能是字母表中的字母、数字、图片，也可能是DNA的化学代码。换句话说，任何可以数字化并能存储在计算机上的东西都可以成为设计空间。因此，我们可以想象所有音乐、艺术、烹饪食谱和建筑设计的图书馆，它们既是有限的，又是没有边界的。

假设我们偶然找到了一个商业人士丢失的笔记本电脑，它可能不小心被落在了机场的X光机上，在那台电脑上我们能找到什么？我们可能会找到一些MP3格式的音乐文件，他或她家人的照片，诸如销售演示、预算提案以及描述各种项目的邮件，甚至是一些被叫做商业计划的东西。

斯密图书馆

你会如何描述一个企业的设计？你可以写下一份有关企业目标的描述，包括它的竞争策略、产品和服务的组织架构、营销和销售计划、生产计划、所需的技术和技能以及经济模式，等等。简而言之，你要写一份商业计划书。[14]

即使是先进的大公司也很少会有单个的、全方位的商业计划。相反，这样一个计划的内容可能分散在组织的各个业务部门的计划、战略陈述、组织结构图、预算案、年度报告、监管文件、注释以及人们的头脑中。同时，许多小企业根本没有正式的商业计划，信息也存在于人们的头脑中。关键是，即使这样一个文档不存在，原则上也有人可以创建它。

继丹尼特的巴别图书馆之后，我们可以设想一个更大的图书馆，它包含了每一个可能的商业计划，这些计划可以被写进500页的书中，或是由几册书构成的一套书中。因此，我们将拥有一层层的商业计划书架，而不是文学书架，这些书架会延伸到宇宙的边缘。事实上，这个包含所有可能的商业计划的图书馆将是巴别图书馆的一个特殊部分。我将这个商业计划子库称为斯密图书馆，这个名字源自亚当·斯密。

如果我们在斯密图书馆的书架上搜寻，就会发现时代广场上一家个人擦鞋店的商业计划，一份完美描述IBM公司现有战略的商业计划，1952年通用电气公司的商业计划，公元前8 500年黎巴嫩一个小麦农场主的商业计划，以及一份亚诺玛米人狩猎队伍的商业计划。斯密图书馆还包含一份生产并销售超纳米神经囊胚管的公司的商业计划书，尽管超纳米神经囊胚管到2023年才能被发明。斯密图书馆描述了人类为谋生所创造或将创造的所有方式。当然，像所有的设计空间一样，图书馆里绝大多数的书都是完全没有意义的商业计划，只有极少数会变成现实。

但这些想象的计划得多具体呢？我们如何能将一份商业计划书和《白鲸》的副本区分开来呢？

我们将使用的测试与用在其他任何概要上的测试相同。斯密图书馆中的一本书是否是有效的商业计划，关键在于商业计划书阅读器是否能依照这个计划组织并创造经济活动。什么是商业计划书阅读器？它是一个管理团队。

假设现在是2006年，我们从斯密图书馆的书架上拿下一本叫做《思科系统2007年商业计划》的文件，把它交给思科系统现有的管理团队。如果团队能够获取该计划中的信息，执行并实现该设计，那么它将是一个有效的商业计划。如果我们为思科管理团队提供许多500页的、文本乱七八糟的书，那么什么都不会发生。如果文本是《白鲸》，他们将可以顺利理解其中的内容，但思科的业务不会发生改变。但是，如果我们给管理层一份文档，它可以转换为思科系统的新设计，那么它就有资格成为思科系统的商业计划和候选设计。当然，图书馆中会有大量潜在的

有效商业计划。一个计划可能主张积极的国际扩张，另一个计划可能主张推动新产品开发，还有一个计划可能建议缩减投资，而其他计划可能是这些策略的组合。一些计划是有利可图的，但大多数不会。还有一些其他的计划提议对思科系统现有的业务逐步进行改造，进而实现荒谬的彻底改变，比如一份建议公司退出计算机通信领域并进入快餐业的计划。

一个反对意见可能是，任何想象中的思科商业计划，即使写了几千页纸，都无法涉及构建思科系统的每个细节。然而，反对意见适用于将概要呈现到设计中的任何过程。通常说来，概要是一种速记形式，是人们试图描述的设计的浓缩。因此，并非运行思科系统所需的每个细节都要在有效的计划中描述。商业计划依赖于读者已有的隐含知识、背景和技能，正如我们的乐高概要假设儿童阅读器在挑选和组合塑料积木上有特定的技能，或者DNA假设有具备必要能力的卵细胞或子宫的存在。因此，一个计划需要有足够的信息才能保证效力，让2006年的思科系统管理团队知道如何处理它。

如同生态系统，商业计划的概要是和它们的阅读器共同进化的。例如，2005年的思科公司商业计划（无论是书面的还是仅仅存在于人们的头脑中的）影响着2006年的管理团队类型，成员需要具备的技能、知识以及经历。反过来，这将定义团队能够阅读并执行的2007年的计划，2007年的计划将影响管理团队的未来发展，进入一个共同进化的循环中。[15]如果我们把一份亚诺玛米人狩猎队伍的商业计划交给思科公司的管理团队，它对雨林部落成员而言是一个完美的计划，但思科团队也能按照它行动。同样，如果把思科的商业计划交给一群亚诺玛米人，他们也无法突然建立起一家互联网路由器公司。

一个经济进化模型

随着进化在斯密图书馆中寻找适合的设计，无论是亚诺玛米人狩猎队伍的设计，IBM公司的设计，还是2023年超纳米神经囊胚管公司的设计，经济都在不断发展。正如进化算法通过干扰囚徒困境图书馆而创造出创新、增长和创造性破坏的模式一样，进化的干扰也能通过斯密图书馆在实体经济中创造这些模式。

如果我们回到第9章所讲的一般进化模型，接下来可以询问变异、选择和复制过程是如何在商业计划的基层运作的。在接下来的4章中，

我们将研究经济中对适合设计的进化搜索过程。我们将看到，当人们不断尝试、探寻和发明新的商业策略和组织设计时，变化就会发生。选择会在经济的多个层面起作用，导致一些商业计划成功，而另一些失败。同样，成功的设计会获得更多的资源并得到广泛的复制。

经济进化并不是单一设计空间进化的结果，而是横跨三个设计空间共同进化的结果。在第1章中，我简要介绍了哥伦比亚大学的理查德·纳尔逊提出的两个概念：物理技术和社会技术。[16]当想到“技术”这个词时，我们通常会想到物理技术。物理技术是将物质、能量和信息转换成对人类有用的设计和过程，例如，将沙子变成玻璃或硅片。社会技术也很重要，但往往不在我们思想的最前沿，人类用它们来组织自己的设计、过程和规则。村庄、军队、矩阵组织、纸币、法治和“准时制”库存管理等都是社会技术。

商业计划至关重要，它能将物理计划和社会计划融合在一个策略中，然后在经济世界表达最终的设计。我们需要将商业计划、物理技术和社会技术视为三个不同的设计空间，每个空间在工作中都有其独特的功能。商业计划的选择往往是基于经济原因作出的，但物理技术和社会技术可能会基于其他目的而进化。许多重要的物理技术被发明出来满足军事、医疗或其他社会需求，或者仅仅是由于科学家和发明家的好奇而发展起来的。同样，许多社会技术，如法治或普及教育，具有重要的经济功能，但最初可能是为其他目的而开发的。进化系统的一个共同特点是，为一个目的发展出的创新可能会被另一个目的所取代，这一过程被称为预适应。[17]

我将要概述的模型将经济进化看作物理技术空间、社会技术空间和商业计划空间共同进化的产物。我们可以把它们看作三个不同的，但又相互关联且共同协作的设计空间。在每一个空间中，进化都在起作用，干预可能的设计，发现并放大有效的设计，摒弃无效的设计，从而创造出了我们在技术、社会和经济世界中看到的秩序。

11

力量二：
物理技术

斯坦利·库布里克（Stanley Kubrick）的电影杰作《2001：太空漫游》，以一个名为“人类黎明”的场景开场：太阳从东非稀树草原永恒的美景中升起，一群类猿生物和清晨一起苏醒。这些生物是早期的原始人，不是猿，也不是人类。原始人进入了一个水坑，这是干旱地区的稀缺资源，在那里，他们碰到了另一群原始人。第一群原始人规模稍大，攻击性更强，在进攻中的组织更有序。这是一场势均力敌的战斗，最终，第一群原始人设法赶走了其他人，得到了水资源。规模、侵略性和社会组织带来了优势，这一结果并不是进化上的意外，同样的故事主角也可以换用其他许多物种来讲述。

在另一个场景中，几十万年后，一个稍大的、更擅长直立行走的原始人在一头野猪的骨头堆中搜寻食物。为了便于抓住树枝，他的手进化得十分灵巧。他紧紧握住猪的大腿骨，然后开始慢慢地摆动。随着那位原始人越来越兴奋地破坏猪骨头，他的大脑中建立了一个关键的联系：他可以用大腿骨作为棍子，砸碎活猪的头骨，就像对待死去的猪的头骨那样——这是一个非常有用的联系。这种发展在进化上有点不寻常，但并非完全独一无二。其他物种，特别是灵长类，可以把发现的物体用作工具。例如卷尾猴可以用石头砸开坚果，黑猩猩可以用木棍把昆虫从树洞里掏出来。

几十万年后，一群原始人围坐在一起，他们把骨头和树枝做成了可辨认的棍棒。一旦我们的祖先真正开始制造工具，而不是简单地找到它们——或者像人类学家说的那样创造人工制品，而不是简单地捡起自然物品，人类就在进化上迈出了远离其他物种的重要一步。

这个场景还展示了一个成年原始人如何教孩子制作一根棍棒，从而通过文化将制作工具的知识传递给后代。教授年青一代和传播文化知识都不是人类独有的，比如母狮子就会教它们的幼崽打猎。根据进化论理论家理查德·道金斯的观点，一旦山雀学会用嘴撬开奶瓶，这种技能就会通过模仿在整个种群中传播开来。[1]但只有人类才能把制造工具和向同龄

人及幼儿传播知识结合起来。正是在这一点上，进化才真正地从我们的身体中跃出，进入我们的社会文化，那也是物理技术诞生的时刻。

库布里克镜头下的原始人将他们新发现的技能运用得很好：狂热地将野猪棒打致死，然后享受由此而来的热量和蛋白质盛宴。但随后，他们又陷入了与一个欠发达的原始部落的另一场对峙中。这一次，挥舞着棍棒的手工品创造者们连续击打着他们的对手。在庆祝活动中，胜利一方的男性领头人抛出他的棍子，那棍子在空中旋转，然后诗意地变成了一个围绕地球旋转的宇宙飞船。库布里克的观点是正确的。一旦工具制造和文化知识的传播发生了最初的突破，它就成了一个进化的跳板，使人类跃入拥有宇宙飞船的现代世界。

经济人的黎明

在这个故事的非好莱坞版本中，这种制造工具的突破可能发生在250万年前。证据显示，人类历史上第一个工具是由能人制造的粗糙的片状石质手斧，它是在坦桑尼亚的奥杜威峡谷中被发现的。大约100万年前，人类进入了直立人时代，开始制造一些复杂的工具——手斧两侧都有细微的敲击痕迹和剥落，还有各种各样的设计。此外，直立人还发现了另一种非常有用的工具——火。在距今260万～100万年前，我们的祖先开始用他们的大脑让物质和能量屈服于他们的意志，并开始与他们的孩子和同龄人分享这些知识。

当然，没有人确切地知道人类的经济活动最初是在哪里、何时以及如何产生的。但在这期间，可能有两个原始人哈利和拉里坐在一起，哈利咕哝着问他的同伴：“你想用那把斧头交换一块肉吗？”[2]哈利这天很幸运，获得了一些额外的肉，而他不是很擅长做斧头，也可能只是不喜欢做。拉里觉得做斧头容易又有趣，但他今天的打猎不顺利，于是拉里回答说：“是的，当然。”最终，他们决定进行物物交换，于是经济诞生了。[3]

经济依赖于两个因素的存在：使人们创造值得交易的产品和服务的物理技术，以及使非亲属在创造和交易这些产品和服务时顺利合作的社会技术。证据表明，这两个因素都存在于能人和直立人的社会中。因此，我们可以说，在那一点上，经济诞生了（尽管长途贸易直到很晚才开始）。[4]

早期原始人制造工具无疑是一个有限但富有逻辑的演绎过程，例

如“如果我用一块更硬的石头敲击，较软的石头就会更容易脱落”，这就导致了实验，例如尝试用不同硬度的石头相互敲击。成功的实验催生了制作工具的经验法则，例如“这种类型的岩石有利于削手斧”。然后，这些归纳出来的技术秘诀就可以传给儿童和同龄人，从而避免每个人都必须“重新发明斧头”。有了这些广泛传播的技术秘诀，人类开始了一个步步为营的物理技术进化过程，因为每一代工具都建立在过去工具的制作基础之上，这使得工具随着时间的推移变得越来越多功能，越来越复杂。我们在考古记录中可以找到这些证据，从这些记录中可以看出，一代代工具随着时间的推移进化得越来越快。

物理技术空间

我们可以在一代又一代的现代技术中看到同样的进化模式，例如汽车从T型车发展到充满微处理器的现代汽车，或者手机从手提箱大小发展到口袋大小。[5]人们还可以看到不同技术之间的关系，它们看起来就像是物种形成——飞机与热气球、飞艇和悬挂式滑翔机有关，它们属于同一种人工飞行制品。我们也可以看着技术走向“灭绝”。例如，在华盛顿特区的中部，人们可以发现19世纪旧运河系统的遗迹，在其全盛时期，这条运河上布满了装载着煤、食物和其他货物的驳船。如今，运河被人们用作慢跑的小径，仍然有几艘古老的驳船被拴在岸边，就像一个物种博物馆里的乳齿象标本。

但是技术进化仅仅是一个我承诺要避免的那种比喻吗？我们通过展示技术进化能够用第9章描述的一般进化模型表示来避开比喻陷阱。一个好的起点是给物理技术下一个更精确的定义：[6]

> 物理技术是为实现一个或多个目标而将物质、能量和信息从一种状态转换为另一种状态的方法和设计。

例如，原始人拉里拿了几块石头（物质），燃烧了一些卡路里（能量）互相敲击它们，创造了一个手斧（设计），从而有了一个能用来砍动物骨头的工具（目标）。或者，一个程序员利用自动售货机中的食物来给他的大脑提供能量，然后电能进入到计算机中（能量），将代码（信息）从一种状态转换为另一种状态，从而创造出人们可以用来娱乐的（目标）视频游戏（设计）。一些物理技术生产了人工制品（例如一个手持斧头或计算机程序），而另一些则提供了服务。例如，人们可以

想象一组物理技术用于发放银行贷款或提供指压按摩。服务中包含有目的的物质、能量和信息转换，例如指压按摩包含将能量（按摩人的卡路里）转换为特定的运动（设计），目的是使接受按摩者的肌肉放松。

一项物理技术并不是事物本身（手斧、软件或按摩），相反，它既是事物的设计，又是制作它的说明和技术。设想一项物理技术的一种方法是思考一个手工品的说明手册或者一项服务是什么样子的。例如，制作石质手斧的手册可能包括成型斧头的图解、所需石头类型的描述以及将它们组合在一起创建预期设计的方法。就像斯密图书馆一样，我们可以为物理技术定义一个模式——一个由自然语言、数学符号、图片、实物演示视频等组成的概要。任何在理论上可以数字化的东西都可以成为概要的一部分。原则上，即使是语言指示和隐性知识也可以被编码到概要中，例如人类学家可以观看拉里用手斧干活，然后将他所见详细转录为文字。

我对物理技术的定义在很多方面与技术的传统经济学概念相似。传统经济学理论包含了将原材料、资本和劳动力转换为产品和服务的“生产功能”，而技术被定义为使转换发生的方法。在传统理论中，技术通常是一个外生的黑箱，而现在我们将把它的结构看作一个进化系统，并试图理解黑箱是如何运作的。

我们接下来可以想象一个比宇宙更广阔的设计空间，这次是所有可能的物理技术图书馆。[7]就像其他图书馆一样，它在任何时间点上都是有限的，它的大小由最长、最复杂的可阅读概要的长度决定，但这个图书馆是没有边界的，可以随着时间的推移扩大或缩小。我们可以想象有近乎无限的蓝图、设计、说明以及秘诀书架，它们能以多种方式对物质、能量和信息进行转换。

就像其他设计空间一样，图书馆中绝大多数可能的物理技术设计都是垃圾。尽管设计在物理上是可以实现的，但大多数都是无意义或者荒谬的。图书馆里到处都是方轮的自行车、底部有洞的水壶和用奶酪做成的锤子，这些都是最清楚易懂的设计。经济上可行的产品和服务设计是整个图书馆中一个无穷小的子集。几乎所有功能清晰的设计都没有任何经济价值，因此，技术进化的任务就是在一大堆无用的草垛中找到可行的物理技术这个针，然后由经济进化寻找具有经济价值的更小的物理技术子集。

物理技术阅读器

和我们讨论过的其他设计空间一样，要使概要合格，一项物理技术必须可以被概要阅读器阅读。因此，我们需要人或者事物从物理技术中提取信息，把它转换成现实世界中的物质、能量和信息模式。也就是说，我们需要有人接受建造房屋的指示，然后把材料变成一座房子。我们的标准并不是说概要含有足够多的信息以至于每个人都能做，而是一个能够阅读计划、制作物品或提供服务的合格阅读器。因此，我们期待有一个合格的房屋建设者团队可以利用物理技术进行房屋建筑设计，或者是一个科学家和技术人员团队可以使用一套药物物理技术来制造雷洛昔芬药片。和以前一样，物理技术的指令不必存在于一个文档中，而是可以分散在各种文档和人们头脑里的知识中。关键原则是，理论上的指示可以以某种形式被记住，并传达给其他阅读器。就像孟德尔图书馆和斯密图书馆一样，概要和阅读器是共同进化的。例如，随着房屋建设者的技能和方法的改变，房屋设计也随之改变；随着房屋设计的改变，建造房屋所需的建设者类型也会改变。

有人可能会反驳说，概要无法传播隐性知识的某些方面。如果知识不能以任何方式被获取，那么它也不能被传播。我们会将这种知识称为艺术，然后把它排除在图书馆之外。为了让进化起作用，知识必须是可传播的，因此它能以某种形式被编码。如果拉里有一些如何制作特别漂亮的手斧的隐性知识，但他不能用任何方式编码或传播这些知识，那么，这些知识将与拉里一起死去，而不会成为未来制作手斧的物理技术进化的一部分。

在人类历史的早期，物理技术只是描述如何制造工具的经验法则，并且最初是通过模仿、嘟哝和身体语言传播的。非语言信号是一种低带宽的编码和传输概要的方式，显然会限制物理技术设计空间的大小和复杂性。此外，嘟哝等方式并不是一种非常可靠的通信方式。错误会不可避免地进入早期的物理技术传输中，但为了使进化正常运作，信息传输必须有一个最低准确性。[8]因此，语言在编码和传输更大、更复杂的物理技术以及更准确地进行编码和传输的能力上是一个巨大的突破。考古记录显示，在旧石器时代晚期，工具制作的创新和多样性发生了戏剧性的突破。那个时代出现了各种各样的材料，鱼钩和缝纫针等新设计开始出现。尽管对语言能力到底形成于什么时候仍有许多争论，但一些研究者认为，工具设计的突然辐射是语言发展的有力证据。[9]

物理技术促进自身发展

人类物理技术最卓越的地方是每一项发明都创造了更多发明的可能性和需求。斯图尔特·考夫曼指出，内燃机的发明促进了汽车的发明，这又促进了充气橡胶轮胎、挡风玻璃雨刷、沥青路面、汽车旅馆、快餐店、收费站和拉斯维加斯的婚礼小教堂的发明。[10]每一项发明都为未来的发明开辟了新的天地。一项发明的成分经常会以新形式被重新使用。正如本书之前讨论过的，一些发明引发了巨大的变化，比如说汽车，而一些发明带来的变化很小（考夫曼认为其规模遵循幂定律）。它们的共同之处在于，所有的发明都有涟漪效应。[11]为什么技术会有这种指数级的自助品质呢？技术如何促进自身的发展呢？

物理技术和其他概要一样，有着模块化的区块构建特征。任何物理技术都可以看作组件和整体结构的编码。[12]一座房子的组件包括房间、管道系统和窗户，其整体设计可以是模拟都铎式建筑。同样，指压按摩也有一些单独的组成以及一个结合它们的设计。

即使是像手斧这样原始的东西也可以用这种方式来思考。手斧基本上是一种两边被另一块石头削出锋利边缘的石头。人们可以把手斧的部件看作使用者握住的石头的一部分，也可以把它看作做成刃口的石头的一部分。建筑本身就是整体设计，在这种情况下，它就相当于有把手的刃口。组件和结构的各种组合定义了设计可能的变化数量。举个例子，有两种不同的把手，一种是圆形把手，另一种是管状把手。刃口的设计可能有两种变化，一种是一边开刃，另一种是两边开刃。我们也可以在结构层面上做一些变化，例如，大一点的或小一点的，花岗岩的或燧石的。定义了这些方面后，我们就可以列出手斧可能的变化形式，例如，“大/花岗石/圆形/双面”“大/燧石/圆形/单面”“大/花岗石/管状/双面”等16种手斧的变体。我们还可以用这种方式描述更复杂的人工制品的变体。例如汽车“V6/1.8升/200匹马力/四门/真皮座椅”或个人计算机“2 GHz奔腾4/128MB内存/80 GB硬盘驱动器/4x CD刻录机”等。[13]

正是这种组件和结构结合的特性使得物理技术比宇宙更广阔，这也意味着创新会使物理设计空间呈指数级展开。我们假设制作手斧有两种方法，一种是单面开刃，一种是两面开刃。后来有一天，原始人哈利偶然发现了一种更好的开刃方法——他使用越来越小的岩石去开刃，使手斧的边缘更锋利。现在我们有三种刀刃类型：单面、双面和细边。有了这个单一的创新，手斧的可能变体数量就从16个增加到了24个。或者，从产品多样性的角度来看，我们可以说，在物理技术空间中发现的潜在库存单位数量增加了8个，或者说增加了50%。同样，每次英特尔在其

微处理器系列中添加新的变体，计算机市场上潜在的库存单位数量也会大幅增加。

结构中的创新也能使库存单位的数量大幅增加。比如，有一天，哈利拿着一把大的管状柄手斧，他没有把它削成一个锋利的刀刃，而是把它做成了稍钝一点的形状。[14]尽管这开始看起来像是一个小组件的改变，最后却成了一种全新的结构——一种可以用来研磨坚果、捣碎谷物及其他有用的东西的石臼。在没有进一步创新的情况下，哈利立刻为石器市场增加了4个新的潜在库存单位："大/花岗岩石臼"、"小/花岗岩石臼"、"大/燧石石臼"和"小/燧石石臼"。创新增加的库存单位数量随着工件的复杂性呈指数级增长。一个只有两个组件，每个组件有两个变体的结构有4个可能的库存单位；一个有着3个组件，每个组件有3个变体的可能结构有27个；一个有着4个组件，每个组件有4个变体的可能结构有256个。结构的创新还可以带来组件的创新。例如，个人计算机结构的发展促进了组件设计的广泛创新。因此，每一项创新都会通过物理技术空间发送或大或小的冲击波，并且可能的库存单位数量会随着每一项发明的出现呈指数级增长。

然而，物理技术空间中可能的库存单位只有无限小的一部分会被考虑到，更不用说被用到产品和服务当中。尽管目前市场上可能有数千种，甚至数万种个人电脑变体，但在全部可能的电脑变体中，只有很小的一部分被制造出来销售。苹果设计生产了一台电脑，它的Power个人电脑处理器是令人惊讶的半透明橙色。戴尔设计生产了一台电脑，它的奔腾处理器是基础的黑色。目前还没有人找到一个很好的理由为计算机设计一个半透明橙色的奔腾处理器，因此，这种变体仍然是一种潜在的、尚未在物理技术空间实现的库存单位。新发明的关键作用不仅在于它们为经济体增加了实际的新库存单位，而且它们还打开了潜在的库存单位空间。每项发明都为更多的发明创造了更大的空间，就像从手斧到库布里克环绕地球的宇宙飞船的跨越。

技术进化中的演绎-探寻法

在第9章，我们讨论过与进化系统相结合的适合度景观是如何大体上联系在一起的，也就是说，既不是完全随机的混乱，又不是一个单峰的完全有序的景观。进化景观通常介于两者之间，比如阿尔卑斯山。有关物理技术空间的一个重要问题是，它的适合度景观也是大致相关的，

因为这将决定进化是否是搜寻它的最佳方式。

一个大致相关的物理技术空间是有直观意义的，这意味着空间中密切相关的设计通常具有相似的适合度。例如，我们预计两台设计相同但火花塞参数略有不同的发动机，可能会有相似但不相同的性能。然而，这种普遍的相关性并不完美，火花塞设计中大多数的微小变化可能不是很重要，但某些差异可能会非常重要，并导致发动机无法工作。因此，尽管与高峰相连的景观倾向于靠近其他高峰，但其本质仍然是崎岖而混杂的，有高原也有瑞士奶酪孔洞。[15]

支持大致相关的物理技术空间的另一条证据实际上是一个存在证明。如前所述，进化是寻找大致相关的空间的高度有效的算法，并且进化过程也倾向于创造大致相关的空间。[16]如果我们能表明人们参与物理技术创新的过程遵循进化搜寻过程，那么物理技术空间很可能是大致相关的。

人们在考虑物理技术创新是一个进化过程时可能会问："进化是一个盲目、随机的过程，而技术创新是由人的理性和意图所引导的。我们怎么使那个圆圈变成正方形呢？"答案是，在进化算法的本质中，没有任何基础可以说明意向性及合理性不能发挥作用，也没有任何基础可以说明这个过程必须是完全随机的。进化论的核心是一个反复实验、选择、放大有效性状的过程。生物进化过程中随机的部分是通过创造选择的多样性来实现的。[17]但即便如此，它也不是完全随机的。突变可能是随机的，有性生物的重组却不是。对配偶的竞争可以确保适合的有机体与其他适合的有机体有更高的配对概率。

唯一的要求是要给算法提供足够多的实验来进行选择。实验必须覆盖足够宽的适合度景观，使算法有微小但值得一搏的机会找到高峰。从算法的角度来看，过程如何实现多样性并不重要。在人类搜索物理技术空间的案例中，进化算法通过"演绎-探寻法"来获得实验的多样性。"演绎-探寻"一词融合了心理学家唐纳德·坎贝尔（Donald Campbell）的理念和赫伯特·西蒙的"有目的的适应"这一概念。[18]基本上，人类在发明行为中会经历两个认知过程，其中之一是他们的演绎能力。简单说来，演绎是《2001：太空漫游》那个场景中展示的原始人建立起棍棒可以用作捕杀武器的思想联结。从复杂的角度来看，演绎是英特尔的工程师使用量子力学理论来计算将电路封装到芯片上的极限。

这个模型的另一个部分——探寻，就是我们展开尝试的地方。正如现代工程使用科学一样，仍然有很多"让我们试试看会发生什么"。[19]历

史上有许多人们意外发明新事物的著名故事。例如，20世纪80年代，3M公司的化学家斯宾斯·西尔弗（Spence Silver）试图制造出更强的黏合剂，却意外地制造出了一种较弱的黏合剂。他的同事阿瑟·富莱（Arthur Fry）随后用这种黏性较弱的胶制作了一个黏性书签，最终变成了现在普遍存在的便利贴。[20]便利贴不仅是借由演绎逻辑发现的，探寻也起了主要作用。正如杜克大学的工程学教授亨利·彼得斯基（Henry Petroski）所说，“常规与失败如影随形”。[21]探寻的本能感觉之下，毫无疑问存在无意识的归纳认知过程、联想思维和类比推理，因此探寻也不是一个真正的随机过程。尽管如此，探寻还是超越了我们仅用演绎法所能想到的，从而扩大了实验的范围。

除了演绎逻辑和归纳探寻外，一种随机的疯狂元素也可以告诉我们人类要尝试什么。1982年的一天，一个33岁的加州卡车司机拉里·沃尔特斯（Larry Walters）去了西尔斯，买了一把草坪椅，在上面绑了45个充氦气的气象气球，带着一个啤酒冷却器和一把气枪，升到了4876.8米的高空中，使整个南加州的空中交通都中断了。[22]他原本打算通过气枪射击气球来着陆，但由于过于紧张，他射击了几个气球后气枪就掉下去了。几个小时后，他缓缓飘回地球表面，随后被电线缠住，使长滩的大部分地区陷入了黑暗。最终，沃尔特斯设法爬到了安全地带。沃尔特斯显然对任何航空理论的演绎试验都不感兴趣，他也没有研究飞行草坪椅的商业潜力——他只是想飞。即使有了所有的科学知识，还是会有一些人尝试去做疯狂的事情。尽管飞行草坪椅还没有成为一项重要的新技术，但这种疯狂的东西有助于保持实验在物理技术空间中的广泛传播，每隔一段时间，某些东西就会起作用。

为了描绘演绎-探寻模型的效果，我们可以想象一个阿尔卑斯山式的适合度景观上的一个点。这个点代表特定物理技术的设计，例如微芯片设计。然后，我们在这一点周围的景观上画出一系列的实验，你的绘图可能类似于图11-1。我们会看到许多触手般的探索之手从当前的设计中发散出来，这些发散物是与演绎探索高度相关的分支。这些分支会进入到理论、科学和过去的经验显示富有成果的景观领域，并避开那些不大有希望取得成果的领域。在微芯片的例子中，这些分支代表的是芯片工程师为改进其设计所做的精心排布和努力。接下来，我们将看到一块探寻实验区从手指一样的树枝周围渗出，这代表的是芯片设计者更大胆、更具实验性及面临更大风险的想法。我们会在景观图中看到一些随机的点，这些代表的是人们的一些“不靠谱”的尝试。如果实验是由随机的变化和杂交驱动的，那么实验在景观上的分布将会呈现不同的形状（见图

11-2）。但关键是，这种扩散足以让进化发挥作用。

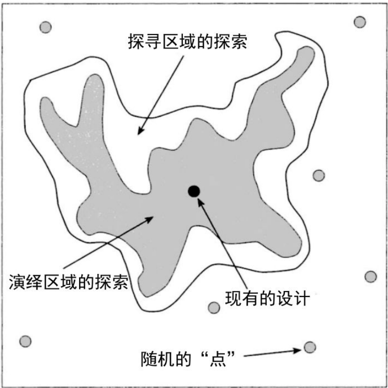


图11-1　景观上的演绎-探寻模型

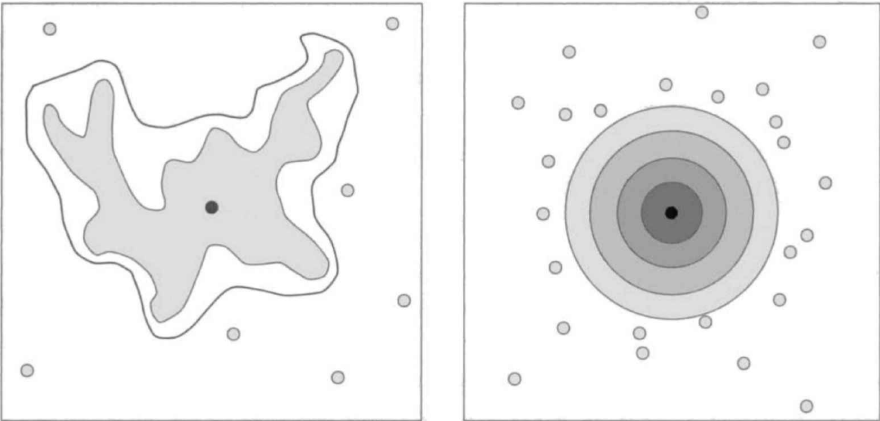


图11-2　演绎-探寻VS. 随机的变异和杂交

尽管演绎-探寻创造出的模式（左）与随机的交异和杂交（右）不同，但两者都可以在景观上传播实验，为进化提供差异的来源。

物理技术景观选择

演绎-探寻为物理技术的变革提供了一个机制，那么，选择是如何进行的呢？什么决定了物理技术景观的适合度呢？

因为有了变化，这个过程不同于生物进化，但仍与一般算法的运行一致。所有的物理技术都是为了实现某种功能，[23]手机的设计是为了促进交流，咖啡杯的设计是为了装咖啡，塑料雪球纪念物的设计是为了提醒我们在蒙特利尔的美好旅程。物理技术选择的基础是其对目的的适用

性，我们会寻找那些比其他物理技术更能满足某种目的的物理技术。“更”意味着它们能更高效地满足目的，而且在经济上比其他选择更节约。因此，物理技术在制作雪球领域的产生和广泛使用，初步证明了许多人发现雪球既能提供娱乐又具有纪念意义。塑料雪球在全球的存在证明，相对而言，它们在竞争环境中比那些纪念咖啡杯、劣质的茶巾和印有“我朋友去了蒙特利尔，而我得到的只有这件糟糕的T恤衫”的T恤衫更有优势。

要想让选择在进化系统中发生，系统就需要有西北大学的乔尔·莫基尔（Joel Mokyr）所说的超凡创造力——比环境所能支持的设计更多的设计，这样才能创造竞争。[24]在生物学中，潜在的有机体设计比任何生态系统所能支持的都要多。技术层面上，可能的物理技术比人们使用的要多，例如，比世界各地的游客所想要的更多的小饰品设计。进化能在物理技术空间中运作，因为物理技术的超凡创造力能使这些技术相互竞争，而且随着时间的推移，人们倾向于选择能有效且高效实现目标的物理技术。

模仿是最真挚的恭维

进化清单上的最后一站是复制。测试进化是否成功的关键是成功的设计在总设计中的频率有所增加。在生物学中，这个过程相当简单。有助于增加适合度的基因比不能增加适合度的基因更有可能传给下一代。因此，适合的基因频率就增加了，但它是怎么服务于技术的呢？不同的物理技术是如何在全部物理技术中复制并增长的呢？

很简单，对人类有益的物理技术就能得以复制。例如，公元前5000年，第一种可以用来制作模型砖的物理技术被发明了。[25]这项非常有用的发明被广泛复制，通过演绎-探寻，用于制砖的各种物理技术变体不断进化。制作砖块的信息被储存在人们的大脑中，被嵌入到人工制品中，并记录在书面描述中。人们观看复制过程，大师教导学徒，人们自己看着砖块寻找线索，最终写下描述，砖块制作的技术迅速传播到世界的每个地方。

当物理技术从一个人的大脑传播到另一个人的大脑，当体现着物理技术的人工制品被复制，当物理技术被刻在石板上、印在书上，并放到网页上时，物理技术得以复制。我们可以认为物理技术具有随着时间变化的“市场份额”，它随着成功物理技术的传播和失败物理技术的消亡而

波动。例如，在维多利亚时代，与砖块有关的物理技术在伦敦达到了鼎盛时期，但在20世纪，它却输给了使用玻璃和钢的物理技术。因此，我们可以说，砖块制作的物理技术在一段时间内享有复制优势，但随着其他竞争物理技术的普及，这一优势就减弱了。

有了物理技术进化的轮廓，我们现在可以简单地看一下它带来的启示。首先，我们将看到大致相关的景观几何形状如何能够帮助解释技术的S形曲线现象；其次，可以研究不同物理技术之间的模块化和相互关系如何解释颠覆性技术的影响；最后，可以研究科学如何改变物理技术的进化过程。

解释技术S形曲线

20世纪80年代，我在麦肯锡公司的同事理查德·福斯特（Richard Foster）提出了技术的自然生命周期理论，并在他的《创新：攻击者的优势》（*Innovation: The Attacker's Advantage*）一书中对其进行了描述。[26]福斯特的理论基于他的观察以及从帆船到微处理器的案例史。他在所研究的案例和数据中发现了一个惊人的一致模式：在一项新技术发展的早期，其性能较差，进展较缓慢。然而，经过一段时间的投资研发以及各种设计的探寻，该技术的性能会突然呈指数级提升。在此期间，从投资到研发的每一美元都会从这项技术的性能中获得巨大的收益。但随着技术的成熟，性能提升曲线开始逐渐放缓，投资回报率开始下降。福斯特将这种模式称为S形曲线，因为投资于改进技术的努力与其性能的关系图呈现出了一个S形（见图11-3）。

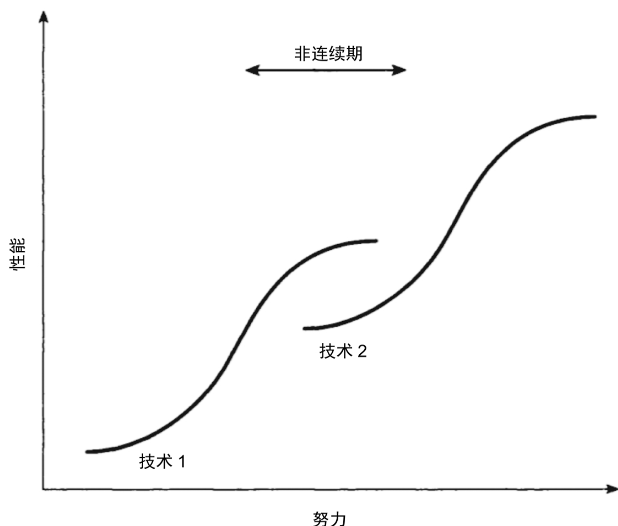


图11-3　技术S形曲线

福斯特理论的另一个部分是，一旦一项技术的投资回报开始减少，企业家就有了寻找新技术的动机。最初的进展趋于平缓，然后达到起飞点，最终，新技术取代了旧技术。福斯特向公司传达的信息是，两条S形曲线之间的非连续期是现任者易受伤害的节点。已获利的参与者倾向于尽可能多地从旧技术中榨取价值。通常，他们会预计过去的性能提升将继续。这些参与者低估了在旧技术的性能不再提升之时，新技术带来的威胁。这就让现任者非常容易受到小型创新者的攻击，而这些创新者都专注于新技术的研发。

如果物理技术空间的适合度景观确实是大致相关的，那么我们将有望看到福斯特观察到的S形曲线。斯图尔特·考夫曼指出，自行车出现早期有大量的实验性设计——大前轮、小后轮以及各种转向装置等。[27]但是，在我们现在熟悉的两个大小相等的车轮结构被设计出来后，基本设计的改进迅速出现，自行车的性能得到了快速提高。然后，随着S形曲线达到顶点，进一步的提升将局限于不同组件的改良，比如更好的齿轮和更轻的框架材料。随着时间的推移，新的结构，如赛车和山地车被发明出来，自行车生产技术的改进随着新S形曲线的出现继续进行。

为了从适合度景观的角度来看这个故事，我们假设有一个特定的点代表第一个自行车设计，以及与它相对应的适合度水平（见图11-4）。这一点周围都是等着被发现的潜在设计，靠近原始点的区域代表自行车设计的微小变化，距离较远的区域代表更彻底的创新。为了提高自行车的适合度，你开始做一些实验，并修改目前的设计，看看是否有一些改

动会改善或降低它的性能。从适合度景观的角度来看，这等同于在附近的区域探索，看更高的山峰和更低的山谷在哪里。

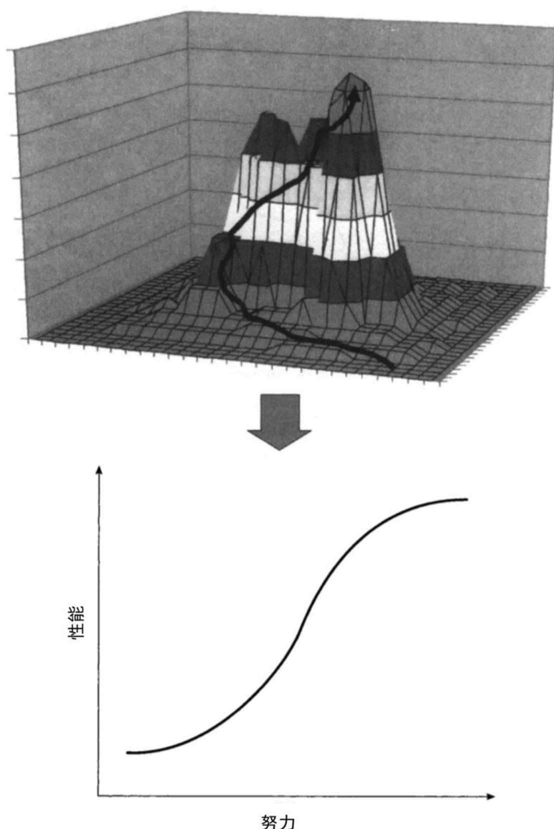


图11-4　大致相关的技术景观与S形曲线

因为你是从发现的第一个自行车设计开始的，还有无数之前从未尝试过的变化。这些变化中有一些会带来更高的适合度，但大部分只会带来更低的适合度，因为建造自行车的坏方法总是比好方法多。一旦你确认了好的设计方向——景观中的高峰，你就开始朝着那个方向行进。不过，刚开始的时候，因为你在用不同的方法修改设计，尝试不同的路径，进展很缓慢。高峰在远处，而你只能在山麓徘徊。但是经过最初的漫步后，你碰上了一个优秀设计（例如两个相同的轮子或后轮驱动），这相当于找到一条通往高地的路线。在一段时间内，自行车在适合度方面的提升是非常迅速的，你每走一步，都会在高度上有显著的提升。最终，你开始接近顶峰，你向上攀爬的速度会减慢。当爬得越来越高时，你会注意到一些重要的事情：每走一步，可能的向上路径数量会减少。在开始的时候，你面前会有许多潜在的方向和未经试验的设计。但是，

一旦你踏上自己的道路，接近顶峰，潜在的向上路径数量就开始减少，变成越来越多的渐进式改进。

如果物理技术空间的适合度景观完全是随机的，我们就不会看到S形曲线了。投资某项特定技术的回报也将是随机的，因为景观上的任何一步探索都会导致适合度的随机变化。如果在另一个极端，物理技术空间的适合度景观是光滑并且完全相关的（像富士山一样），那么将只有一个最好的自行车设计。我们不会争论闪电04型自行车是否会比戴尔R500的齿轮头更好，也不会经历从一条S形曲线跳到另一条S形曲线的不连续阶段。在一个完全相关的空间里，一旦设计达到顶峰，就没有别的地方可去了。相反，S形曲线是粗相关的适合度景观的几何学自然结果。

颠覆性技术

在《创新者的窘境》（*The Innovater’ s Dilemma*）一书中，哈佛大学商学院的教授克莱顿·克里斯滕森（Clayton Christensen）问，为什么从一个S形曲线跳到另一个S形曲线往往会成为成功的大公司的主要绊脚石。克里斯滕森研究计算机硬盘产业的演变时发现，技术上的微小创新往往具有较大的颠覆性。

克里斯滕森声称，一项技术是否有颠覆性与这项科技进步本身的价值关系不大，更多地取决于它对S形曲线的具体影响。如果一项技术将产品性能提升到了现有的S形曲线上，即使速度非常快，该技术都倾向于保持现有参与者的能力。然而，当一项技术需要一条新的S形曲线时，特别是当它开始时的性价比比当前的技术差时，新技术往往能颠覆行业结构。这是因为现有成功公司中的任何一位总经理都很难证明在一项低性价比的技术上投资是合理的，至少在最初是这样。最初的3.5英寸磁盘速度慢，价格昂贵，存储空间小，而一家领先的5英寸磁盘公司的总经理会问：“我们的客户为什么想要那种磁盘？”当一项新技术开始向新的S形曲线爬升时，旧技术的掌握者们就陷入了一场奋勇直追的保卫战中，而他们通常会输掉这场战争。

这就是克里斯滕森所说的“创新者的困境”，它赋予了大致相关的景观下的物理技术空间以合理性。随着时间的推移，硬盘设计从14英寸缩小到8.5英寸，它的挑战不仅在于使同一个产品变小。相反，它需要许多组件的重大改变，新的制造技术以及用电子部件替换许多机械部件。虽

然表面上看这些变化可能是渐进的，但从适合度景观的角度来看，这是一个巨大的飞跃。为了方便起见，我们一直在三维可视化适合度景观，但设计的模块化意味着它们实际上是多维的。如果有人改变了设计中的5个组件，那么它就跨越了景观的5个维度。适合度景观中的“距离”既指改变的维度数，又指每个维度上的变化量。在许多维度上发生的变化，如硬盘大小，可以算作景观上的一个巨大变化。

从定义上讲，结构创新需要一下子改变很多东西，无论是从帆船到轮船，还是从赛车到山地自行车。在对半导体设备行业的详细研究中，麻省理工学院的丽贝卡·亨德森（Rebecca Henderson）和哈佛大学商学院的金·克拉克（Kim Clark）表示，结构创新往往比单个组件的创新更能颠覆行业结构。[28]通过将创新视为对适合度景观的搜索，我们可以看到福斯特、克里斯滕森、亨德森和克拉克的作品都贯穿着一条主线：现有的成功公司很难在物理技术景观中跳得很远。当你身处周围山峰的最高峰时，导致适合度降低的路线比适合度提升的路线多很多，并且选择一个新的结构看起来将面临很大的风险。从一个站在低谷的创业者或者新进入者的角度出发，整个景观中有很多向上的路，也有很多新的无人涉足的山峰可以探索。创业谷中大多数向上的尝试都会终结在没有出路的峡谷或者令人失望的低矮山丘上。但随着有足够多的探索者不断探索，最终还是会有人找到一条有吸引力的上升路线。[29]

科学革命：重新编程进化

物理技术进化的一个关键特征是，与实验性的探寻或随机搜索相比，演绎探索更有可能将人们引向高适合度的山峰。虽然远不完美，但人类的演绎对于预见行为的可能后果，以及进行思想模拟，都是非常有帮助的。在人类99.9%的历史中，我们的演绎能力起作用的范围相对有限，演绎-探寻模型中探寻比演绎的成分要多得多。人类拥有许多能起作用的物理技术，但不知道它们为什么起作用，也不知道它们是如何起作用的。尽管有着种种限制，但人类及其祖先在250万年的时间里还是创造了一系列缓慢增长的发明。

正如我们在第1章中讨论的，在1750年，人类历史上发生了一些不同寻常的事情：物理技术空间以及它带来的库存单位数在经济领域所占的比重迅速增加。科学革命始于1500年左右，意大利文艺复兴时期古典知识的复兴重新点燃了人们对自然现象的兴趣。16世纪，达·芬奇和哥白

尼等人推动了这种复兴。然而，在许多方面看来，达·芬奇和哥白尼更像是探寻者，而不是科学家。17世纪，弗朗西斯·培根阐述了科学的方法，伽利略确立了实验的作用，他们为艾萨克·牛顿、罗伯特·博伊尔以及17世纪后期其他人的巨大进步奠定了基础。自那以后，科学就走上了指数级发展曲线，到今天仍在继续。[30]

科学的作用本质上是大幅地提升演绎洞察力的命中率，因此，演绎突然开始在演绎-探寻中发挥更大的作用。演绎科学理论并不能完全取代探寻，任何一个工程师都会告诉你，理论上有效的东西在实践中并不总是有效的。但平衡倾向可以使二者有效结合，从而使进化发现物理技术景观中的高峰的速度大幅提升。

技术进化不仅仅是一个比喻，它是人类在无限的物理技术空间中进行演绎-探寻的结果。在这种基础上，我们说分化、选择和复制等过程的性质不同于生物学，但它仍然是一个进化过程。这意味着物理技术进化遵循着同样适用于其他进化系统的一般规律，物理技术进化表现出与其他进化系统相同的行为，例如，创新推动进一步创新的趋势，以及技术变化的间断均衡性质。但物理技术的进化有一个人类系统独有的特点——我们可以重新编程进化搜索算法。科学使人类能够快速探索物理技术空间，随之而来的工业革命和信息革命改变了社会和地球。从“人类的黎明”到绕着轨道飞行的宇宙飞船，而我们现在需要回到原始人朋友们的身边，因为物理技术只是故事的一半。

12

力量三：社会技术

2002年，国际经济研究所的威廉·伊斯特利（William Easterly）和明尼苏达大学的罗斯·莱文（Ross Levine）对72个富裕国家和贫穷国家进行了详细的研究，试图找到使一个国家比另一个国家富有的原因。[1]他们假设，国家财富的主要决定因素包括自然资源、政府政策以及一个国家的物理技术的发展程度等。伊斯特利和莱文发现，虽然这些因素在某种程度上都很重要，但最重要的因素是一个国家的社会技术状况。法治、产权、运行良好的银行体系、经济透明度、清廉程度以及其他社会和制度因素比其他因素对国家经济的发展水平的影响更大。即使是资源很少、政府能力较弱的国家，如果拥有强大、发达的社会技术，也能取得相当好的成绩。而没有一个社会技术落后的国家能有好的表现，不管他们有多么丰富的资源，也不管他们的宏观经济政策多么有序。

社会技术不仅会影响一个国家的表现，而且也能解释行业和公司在细微层次上的表现差异。20世纪90年代末，经济学家开始注意到美国经济的快速增长。起初，研究人员试图通过物理技术寻找解释。在过去的20多年里，美国在提升计算能力方面的投资巨大，一个主要的假设是经济终于见证了这份投资的回报。

然而，麦肯锡公司全球研究所的同事们对此持怀疑态度，他们深入探究了标题下的生产力数据。[2]结果表明，生产力的真正驱动力是企业如何进行自我组织与管理，即社会技术中的创新。

麦肯锡公司深入研究的行业之一是零售业，尤其是沃尔玛公司对整体行业生产力的影响。20世纪80年代和90年代初，沃尔玛公司在大型商店和高效物流系统方面的创新使其生产能力比竞争对手高了40%。这一挑战反过来迫使其竞争对手模仿沃尔玛的组织创新，并在20世纪90年代末将生产力提高了28%。与此同时，沃尔玛公司继续将自身的生产力提高了22%。仅零售业这一社会技术创新竞赛就占美国这段时期整体生产力提升的近1/4。其他5个领域中类似的社会技术创新竞赛构成了其余的增长。计算机在这个过程中扮演着至关重要的角色，没有计算机，沃尔

玛公司复杂的物流工序将不可能实现。但是计算机技术起到的是促进作用，而不是主要作用，是组织和过程中的创新带来了生产力的巨大提升。

让我们变得有组织

在前一章中，我把物理技术定义为为实现一个或多个目标而将物质、能量和信息从一种状态转换为另一种状态的方法和设计，[3]而社会技术是为实现一个或多个目标把人们组织在一起的方法和设计。

一群人可能聚到一起开创一家公司，创立一个宗教，或者创建一个周五晚上的保龄球联盟。这种组织行为总是在追求一个目标，无论是利润、精神启迪，还是乐趣。正如物理技术是为满足人类需求而在物理领域创造秩序的方法一样，社会技术也是为满足人类需求而在社会领域创造秩序的方法。

社会技术这个词与经济学家使用的一个术语十分相近：机构。诺贝尔奖得主道格拉斯·诺斯将机构定义为“社会中的游戏规则”。[4]机构是组织的一个组成部分，但我对社会技术的定义更广泛一些，囊括其他组成部分，如结构、角色、过程以及文化规范。社会技术包括组织所需的所有要素。一支足球队的社会技术不仅包括比赛规则，还包括守门员的职责、球队的文化，以及球队的战术安排，比如是在前面派出三名前锋，还是在后面派出两名前锋和一名中场球员。虽然足球队的社会技术包括对该队组织方法的完整描述，但不包括该队使用的战略。因此，诸如“左路进攻”或“专注于短传”之类的陈述将不包括在内。在经济背景下，这种战略是商业计划的一部分。

社会技术如何进化

基于社会技术的定义，我们可以为它们构建一个理论设计空间，一个所有可能出现的社会技术的图书馆。建构将遵循和构想物理设计空间时相同的路径。存在于社会技术图书馆里的是为特定的设计和指令创建社会结构的概要，比如为组织一个亚诺玛米狩猎队写一份指示，描述通用电气公司的组织结构，阐述欧洲银行的规章制度。这些指示集可能包括自然语言文本、图表以及描述组织结构、角色、决策过程、规章制度、激励机制、行为准则等的表格。像之前一样，我们可以想象这些社

会技术被收录在500页的多册套装书中——一个比宇宙广阔得多的所有可能社会技术的图书馆。

就像物理技术和商业计划那样，在现实世界中，一些社会技术是以书面形式存在的，但许多社会技术只存在于人们的头脑中。社会技术并不一定要写下来，但是原则上它们要能够被充分地记录下来，进而让一个合格的阅读器能够遵循它们完成设计。一个亚诺玛米狩猎者能够理解为狩猎团队写的概要，一个通用电气公司的总经理能够理解通用电气的组织描述，一个经验丰富的欧盟官员可以理解银行监管结构。

与我们讨论过的其他设计空间一样，社会技术图书馆具有三个重要属性。第一，如同它的表亲物理技术一样，社会技术的设计空间是自我供给并以指数级展开的。[5]每个社会技术突破都可以为接下来的一系列突破创造出更多的空间：货币的发明催生了会计职位，从而带来了股份制公司的发明，以及股票市场的出现等。

第二，社会技术有着模块化的区块构建能力。例如，一家大型跨国公司的组织设计是一系列模块的集合，其中包括组织其业务部门的设计、会计和控制系统的设计、董事会结构的设计以及行为规范的设计。[6]

第三，与社会技术设计空间相结合的适合度景观很可能是大致相关的。社会技术设计上的微小变化会带来相对适合度的微小变化，有时微小变化也可能导致社会技术不可行或者更好。因此，社会技术适合度景观有着与阿尔卑斯山相似的大致形状，带着平滑的点、瑞士奶酪孔洞和通往高地的入口。从这一假设可以得出的预测是，就像我们在物理空间观察到的S形曲线和颠覆性技术那样，我们期待在社会技术空间中看到它们的对应物。历史似乎证明了这一点，[7]例如，从狩猎采集到定居农业的跨越可以被视为人类经济组织在S形曲线上的一个重大转变。[8]同样，亨利·福特于1914年发明了一种组织制造的全新方法——生产线，它是改变早期汽车行业结构的极具颠覆性的社会技术，同时也改变了其他许多行业。[9]

社会技术空间中的演绎-探寻

如果社会技术的适合度景观确实是大致相关的，那么更进一步的启示是，搜索它的高效方法就是我们的冠军搜索算法——进化。就像用演绎-探寻搜索物理技术空间那样，人们也用演绎-探寻搜索着社会技术空间中合适的社会技术。例如，当亨利·福特和他的团队发展生产线时，他

们不仅仅是在纸上对其进行了演绎推理，[10]也不是仅仅尝试了随机试验。相反，他们两样都做了一些。福特被“生产一辆大众能承担得起的汽车”这一愿望所激发，要做到这一点，他需要在制造过程中减少熟练工匠的数量，将更多的工作交给不熟练、工资低廉的工人。福特熟悉美国斯普林菲尔德军械库在制造过程中使用标准化零件方面的进展，尽管他和他的团队没有读过太多的理论经济学，但他们都熟悉专业化分工能带来好处的观点。借助一系列演绎假设，福特在1908年至1912年间开始在工厂中进行实验。经过4年的改进和完善，他在1913年发现了一个关键点，那就是汽车本身应该沿着生产线移动，而不是让工人移动，到了1914年，福特建立了一条移动装配线。

社会技术中探寻所占的比重也相对较高。相比于建造喷气式飞机或研发一种新的心脏病药物，重新设计一家公司的组织架构或者建立中央银行系统等活动中的艺术成分远远超过科学成分。因此，社会技术空间的适合度景观探索模式中，较少有直接的演绎，更多的是试验和试错。复杂经济学的一个承诺是，随着时间的推移，它将把社会技术中的艺术科学边界进一步推向科学。尽管演绎在社会技术空间中的作用较小，但使用演绎-探寻来寻找适合的社会技术仍然是一个进化过程。人们会用不同的社会技术进行实验，然后随着时间的推移，成功的设计往往会持续下去，而不太成功的设计会逐渐消失。成功的设计往往会被复制，因为可以吸引更多资源和传播更广而被放大。例如，福特对移动装配线的创新在制造业中迅速传播，取代了其他社会技术，直至今天仍然是标准做法。[11]

物理技术与社会技术之间是紧密联系的。当人们穿越物理技术的适合度景观时，他们在社会技术的景观上会引起地震和其他巨变，反之亦然。诸如牛拉犁这样的物理技术进步只能发生在农业的社会技术创新之后。同样，现代的许多管理创新在很大程度上都依赖于计算机和通信技术的进步。农业革命、工业革命和信息革命都可以被看作共同进化的旋转木马，物理技术的进步会带来社会技术的新形式，而社会技术的进步反过来又促进了物理技术的进一步发展，如此循环往复。

竞合

是什么驱使着人类在社会技术空间中进行演绎-探寻搜索的？是什么在促使我们不断地寻找新的更好的组织方式？答案在于非零和博弈的魔

力。

在第10章中，我解释了零和博弈，即一个人的收益是另一个人的损失，而非零和博弈是指两个人可以通过合作获得更大的收益。非零和博弈的合作有一个1＋1＝3的逻辑，如果你帮我一把，我就帮你一把，我们可以一起做一些自己无法单独完成，但对两个人都有益的事情。非零和博弈是生物进化中广泛应用的生存技巧之一，比如狗会成群结队地捕猎，白蚁会共同筑巢，鱼儿成群游动，而大多数灵长类动物都过着群居生活。

尽管非零和博弈中的合作能带来实质性的好处，但就像囚徒困境向我们展示的那样，为了更大的利益进行合作以及追求狭隘的自我利益之间往往存在着紧张关系。[12]在发人深省的《非零》（*Non Zero*）一书中，记者兼科学作家罗伯特·赖特（Robert Wright）说，人类历史上的许多事都可以看作合作与个人利益这一核心冲突的结果。[13]赖特称，社会逐步走向复杂的过程，从简单的狩猎采集部落到有组织的村庄，再到民族国家和全球合作，都是人类开创新方法，不断扩展合作规模，改进方法参与更加复杂有利的非零和博弈的结果。他指出，在资源有限的世界中，合作面临着竞争压力。随着时间的推移，能够更好地进行组织的社会将在社会、经济和军事上主导那些在创造合作结构方面不太成功的社会，因此，推动社会创新的是竞合。

用我们的语言改写赖特的观点，我们可以将对社会技术适合度景观的演绎-探寻搜索看作能够使人们捕捉非零和博弈利益的社会技术的追求。社会技术的适合度依赖于三个因素：第一，社会技术必须有提供非零和博弈的潜力；第二，它必须提供使人们有参与博弈的动机的收益分配方法；第三，社会技术必须有应对缺陷的机制。

非零和的魔力

在非零和博弈中，1＋1＝3的魔力有4个基本根源。这4个根源长期以来在传统经济学理论中都非常有名。第一个是分工，亚当·斯密在两个多世纪以前就指出了分工的好处。如果两个人的技能略有不同，那么每个人都可以通过专注于自己最擅长的事情，并进行交易来获得共同利益。如果拉里是一个杰出的猎人，而哈利是一个优秀的斧头制造者，那么拉里最好去参加围猎，而不是徒劳地敲打石头，哈利也最好去制作斧头。

第二个是人的异质性。人们的不同需求和品位创造了互惠贸易的机会。查尔斯·达尔文在乘坐贝格尔号航行时，在与火地岛印第安人交流的过程中观察到了这种贸易的好处：“双方都感到惊讶，目瞪口呆地看着对方。我们同情他们，因为他们用鱼和螃蟹交换破布；他们则在嘲笑用如此绚丽的装饰品交换一顿晚餐的愚蠢人类。”[14]

第三个是扩大回报规模的好处。例如，一个猎人可能会在几个小时的狩猎中投入500千卡的能量，并有20%的机会杀死一只有2 500千卡热量的动物。因此，他的预期回报等于500千卡，刚刚够本。如果他和另外两个人组成了一个狩猎队。三人仍然每人投入500千卡，但他们捕获一只动物的概率上升到了90%。因此，预期值变成了每人750千卡（$2500\times90\%\div3$）。仅仅是加入这个团队，猎人就额外获得了250千卡的“利润”，并且大大降低了自身的风险。

第四个是合作有助于削减不确定性。如果一支狩猎队有一天打猎很成功，另一支不成功，成功的队伍可以把猎物和失败的队伍分享，前提是当情况逆转时，另一支队伍也会这么做。因此，合作是降低风险的一个好方法。如果你只能靠自己并且运气不好，你就会挨饿。但是如果你在一个合作的团队里，你的同伴就可以帮助你渡过难关，直到你有能力偿还他们。

分配战利品

这4个非零和收益的根源可以在不同背景下相互组合，从而创造出无数种促使人们为了共同利益而合作的方法。但要想让人们产生合作的动机，他们就必须分得一份战利品。因此，如何分配合作的收益是一个至关重要的问题。如果收益以错误的方式分配，那么合作就会瓦解，非零和收益也会蒸发。

约翰·纳什（John Nash）1950年在一篇题为《交易问题》的绝妙文章中第一次对合作中的回报分配进行了评论。[15]在文章中，纳什问了一个简单的问题，即两个交易者如何在收益方面达成一致？哈利会给拉里多少肉来换手斧？尽管听起来很简单，但实际上这个问题困扰了几代经济学家。纳什的解决方案提出，两个或两个以上的交易者如何分割交易所得收益取决于每一个交易者对交易收益的重视程度以及双方的选择。每个人都在寻找对自己而言最优的交易，他们会假设其他人也在寻找最优的交易。在知道对方行动的前提下，双方都没有改变立场的动机就能

达成交易。

交易双方都愿意进行交易的点被叫做纳什均衡，因此，当哈利和拉里对斧头进行讨价还价时，他们最终找到了一个双方都乐于交易的点。进行交易对两者而言都比不进行交易时的情况要好，因此他们获得了合作的非零和收益。

但是纳什均衡的存在并不能保证一个快乐的、合作的结果。在单轮囚徒困境博弈中，纳什均衡是避免两个囚徒互相告密而坐牢的方案，因为如果你是一个嫌疑人，你不知道你的同伴是会坦白交代还是保持沉默，你最好坦白。因为你的同伙也面临同样的选择，你可以假设他也坦白了。纳什定理告诉我们，要想让非零和博弈指向合作，就需要精心组织收益，使每个人都倾向于选择合作，或者参与者需要某种机制来协调他们的反应。假设嫌疑人有一个黑帮老大，他会杀了任何作证的人，而在出狱后奖励那些保持沉默的人。这将改变收益结构，并将纳什均衡移动到一个双方都不开口然后获释的点。同样，如果嫌疑人被允许交流，并且知道警方提供给另一方的条件是什么，他们就可以协商回答，避免背叛陷阱。然而，最后的结果仍有可能是一名嫌疑人因为利益出卖了另一名嫌疑人。毕竟，小偷之间没有信誉可言，这就指向了社会技术适合度中的一个关键因素：要想让社会技术适合，我们必须制定机制来惩罚那些不友善的参与者。

欺骗者（基本）从不会赢，而赢家（基本）从不欺骗

欺骗的动机意味着合作本身就很难达成，即使达成了也可能是不稳定的。一个猎人可能比他的朋友跑得慢些，只消耗400千卡，但是，只要朋友没有注意到他的投机取巧，他仍然可以得到750千卡的食物。同样，我用来换你那把漂亮的手斧的肉可能又老又硬，或者我只给你5盎司，而不是约定的6盎司。进化生物学的自私逻辑认为，如果骗子骗了人并逃脱了惩罚，他们就提高了将欺骗基因传给后代的可能性，因此，作弊提供了进化优势。

如果生物选择使得欺骗基因比受骗基因更有优势，那么合作如何在群体中立足呢？答案是，合作带来的收益非常强大，以至于合作基因比欺骗基因更有优势，但这种情况只有在这些基因并不天真，且确保欺骗者会受到惩罚时才会如此。[16]为了生存，合作基因需要一些复杂的防御

机制。这里可以回想一下我们之前对林格伦模型的讨论。在囚徒困境博弈中，如果玩家只玩一轮，两个玩家都有相互告密的动机。然而，当游戏反复进行，没有人知道何时结束时，互动就变得更加复杂了。进化出的稳定而成功的策略通常遵循着这样的逻辑："我会从相互合作的假设开始，如果你欺骗我，我不仅会拒绝合作，而且会惩罚你，即使损失我的近期利益也在所不惜。然而，过了一段时间，我可能会原谅你，并再次尝试合作，因为你之前选择欺骗可能是因为沟通错误，或者你可能已经改变了你的行为方式。然而，如果你再次选择欺骗，我再次原谅你的可能性就会降低，惩罚也会变得更加可怕。"[17]

正如进化在林格伦的计算机模型中产生了这种逻辑一样，进化也在人类祖先的头脑和本能中"植入"了这种逻辑。在前文中，我们讨论了一系列"最后通牒博弈"实验，研究人员给了两个受试者一笔钱，让其中一个人决定如何分配，如果另一个人认为分配是公平的，那么他们可以各自保留自己的份额。但如果另一个人拒绝了分配方案，双方都将一无所获。这些实验的结果令人震惊。从经济的角度看，人们应该接受任何分配方案，因为有钱总比没钱好。然而，在一次又一次的测试中，受试者都拒绝了他们认为不公平的提议。实验结果在世界各地的文化中都是一致的，包括狩猎采集文化。其他游戏和实验都证实了人类合作互惠行为一致且根深蒂固的本质。[18]进化促使我们在朝着一个方向发展——天然倾向于合作，以获取非零和收益。进化也使得我们对欺骗非常敏感，期望公平，并惩罚那些我们认为越界的人。事实上，我们已将进化编程到了智能软件中，复杂、直观的"纳什均衡发现者"和"公平检测器"使人们形成了合理稳定，惩罚搭便车者和防御欺骗者攻击的联盟。[19]

然而，人类的互惠设定并不是固定不变的，它可以随着环境的改变而改变。当我们处在一个大部分的经验都是与他人互惠合作，社会规范向我们传递他人可信赖的信号的环境中，那么我们就会倾向于合作。当互惠设定遇到背叛或欺骗的例子时，它也会更加惊讶和宽容。这就像我们的大脑对周围的人群进行了统计，如果人们通常是合作的，那么在遇到一个骗子时，我们会倾向于假设这个人的行为可能是错误或误解的结果。相比之下，在一个低合作、高欺骗的环境中，社会规范不支持合作，合作设定会使我们倾向于保持怀疑。我们对欺骗最初的信号反应激烈，如果有可能的话也只会逐步原谅，而且可能会拒绝合作，直到对方率先给出合作的信号。

互惠规范的局部调整可以在人类层面产生非常复杂的动力。如果欺

骗达到临界点，高合作型社会就会见证合作的崩塌，低合作型社会就会陷入不合作、经济贫困的死胡同；当来自不同合作传统的人们相遇时，就会导致误解和混乱。[20]组织中也会出现同样的问题，正如我们将在第16章中讨论的那样，公司中高度合作、高度信任的文化往往会带来更高的经济效益，而当进化出不同信任参数的群体突然相遇时，公司之间的合并可能会遇到重大问题。

能更好地利用非零和收益，找到分配这些收益的纳什均衡，管理背叛问题的社会技术将会比那些无法做到这些的社会技术在适合度景观上获得更高的位置。当人们为了寻找适合的社会技术在景观上进行演绎-探寻时，人类已经发展出复杂精密的社会结构来解决这些问题。

从家庭单位到商业单位

社会技术的进化之旅始于我们与近亲进行密切合作的基因偏好。[21]毕竟，家庭成员与我们共享一些基因，帮助他们有助于提升这些基因传给下一代的概率。人类祖先最早的社会合作结构就是家庭，原始人的家庭习惯介于“放荡”的黑猩猩和一夫一妻的猿类之间。在大多数狩猎采集部落中，有些有能力的男性可以娶多个妻子，男性地位越高，他的伴侣越多。[22]早期人类倾向于选择一夫多妻制，但与黑猩猩不同的是，男性通常会陪伴并且投资他们的配偶和后代，创造相对稳定的家庭单位。有些社会从未超越这种最基本的社会结构。例如，罗伯特·赖特注意到，直到最近，爱斯基摩人和印第安人都是围绕家庭单位组织起来的，几乎很少有超越家庭的社会结构。[23]

朝着社会结构阶梯向上攀登的第一步是合作狩猎队。狩猎队的基本热量逻辑如此吸引人——有更多东西可吃，以至于当今大多数狩猎采集社会和早期社会都拥有这种合作形式。然而，早期社会的合作狩猎队规模相对较小，主要由亲属或近亲属组成。

随着定居农业的出现，社会合作出现了巨大的变革。大约在11 000年以前，世界各地的人们独立发现了培育作物的物理技术。[24]卡路里增加和风险减少使定居变得更加稳定，人类群体的规模显著增长，这意味着合作开始延伸到由家庭成员构成的宗族之外，这反过来又开启了一系列新的非零和博弈。[25]合作团体可以利用规模经济，例如建造大型、复杂的住所和其他建筑，从劳动分工中获益，以及与遥远的部落成员进行交易。

所有这些合作中的创新带来了一个新问题：该如何分配由此产生的财富呢？灵长类动物的性别等级制度为此提供了一个答案，罗伯特·赖特和其他人把它称作“大男人社会”。[26]就像其他许多物种一样，在灵长类动物中，雄性动物会通过竞争俘获雌性。身体强壮、聪明、好斗的雄性会战胜较弱的雄性，从而把身体强壮、聪明、好斗的基因传给后代。在一夫多妻制的社会，如早期人类，男性越占优势，配偶越多，相应的后代也越多。这就在实力不同的男性群体中形成了等级制度。随着早期人类群体形成社会，这种性别等级很自然地转变成社会经济地位等级，以分配合作产生的利益。当然，这两种阶级制度是同一个硬币的两面：在古代社会，性别优势往往与经济上的成功相伴出现，而经济上的贫富与性别地位是对应的，因为富有的男性可以为配偶和后代提供更多资源。[27]从古希腊神话到现代小报，人类最喜欢的三个话题是性、金钱和地位，这都与我们的祖先有关。

如前所述，等级制度不太流行，没有人喜欢独裁者。企业都在力争使自己的组织架构扁平化，而商业领导应抵挡作为雄性或雌性领袖在组织中炫耀的诱惑。然而，网络理论表明，等级制度在任何信息处理系统中都起着至关重要的作用，无论是计算机芯片、互联网、人脑，还是经济。要利用分工和规模经济带来的好处，需要有人将任务划分开来，协调人们的工作，把事情组合在一起，然后分配利益。在早期社会中，这个人通常是在性别上占主导地位的男性。赖特描述了美国早期西北部沿海部落中的一个“大男人”的角色：

> 总负责人是政治领袖，他效忠于一个宗族或是一个村庄。他负责策划编织渔网、建造鱼窖，确保一些村民专门从事某些工作，比如制造其他村民可以使用的独木舟。为了支付所有这些费用，他会从猎人的猎物中拿走1/5，甚至1/2。这些收入中的一部分将以宴会的形式返还给人们……
>
> 不用说，那个“大男人”会从顶端捞取一些利益。他住的房子比一般人好，连衣柜也比一般人好。[28]

我们可以看到这和现代的首席执行官和政治家十分相似。

人类一旦发明了等级制度，走向一级套一级的嵌套结构就十分容易了。我们可以想象这个进程：在某个时间点上，一个成功的大男人掌管着一个不断扩大的村庄，他没时间看管负责编织渔网的人，因此他任命他的弟弟或者最好的朋友来管理。就这样，商业团体诞生了。这个大男

人让他的弟弟或朋友向他汇报，又有更低级别的管理者向他的弟弟或朋友汇报。等级制度有助于分工以及信息处理，它普遍存在于人类的社会结构中，从狩猎采集部落到社区保龄球联盟，再到大型公司。[29]

和平、爱和理解

尽管社会和经济等级对信息处理颇为有益，但这样的结构本身是不稳定的。人们会不停地为获得更高的位置展开竞争，大男人失去权力或者死亡后又会不可避免地接连发生战争。组织混乱的代价高昂，而稳定却有许多好处。稳定的组织有能力随着时间的推移积累知识和技能，参与非零和博弈，获取长期回报，为参与者提供更大的确定性，从而以更低的成本吸引人们进行合作。在灵长类部落中，等级管理的方法相当简单：一个强壮、聪明、好斗的男性统治着某个群体，直到另一个更强壮、聪明、好斗的男性在一场暴力政变中将他推翻。这种等级管理方法至今仍被独裁政权、有组织的犯罪集团和其他不受欢迎的社会组织所采用。

随着农业和更大、更长久的定居点出现，人们发展出一系列在等级制度中移交权力的创新方法，避免了代价高昂的暴力。这些创新包括长子继承制和由“长者”挑选领袖。而近代社会的创新包括民主选举以及股份制公司。有趣的是，尽管这些过程在现代都有一定的礼仪，但武力威胁总是潜伏在表面的和平之下。如果一位美国总统在选举失败后拒绝离开白宫，一些戴着墨镜的大块头大概就会把他或她赶走。管理等级结构变化的社会技术对于保持组织的稳定和持续发展来说至关重要。

需要管理的不仅是等级内部的竞争，来自不同等级的成员间的竞争也会带来威胁和机遇。人类需要能使完全陌生的人开展合作的社会技术。陌生人遇到的第一个问题是他们不知道是否可以信任对方。人们表现出合作及互惠行为的规范参数可能非常不同，一方可能在交易中利用另一方。因此，人们需要社会技术来确定除了他们的直系亲属外，哪些人可以信任，哪些人不可以信任。第一个延伸到村庄之外的社会技术无疑是部落身份。通过识别“自己的人”和“外面的人”，你可以高效地找到那些你更愿意与之遵守相同的社会规范的人。而且，由于你在一个部落中的交流可能会随着时间的推移而重复，你就不太可能遇到敲诈或严重的误解。

从历史上来看，交易网络首先倾向于在部落、民族和宗教团体中强

势发展。比如公元250年到900年间覆盖墨西哥、危地马拉和伯利兹大片地区的玛雅交易网；在过去千年中从北非延伸到中东和中亚的穆斯林交易网；以及主导华尔街公司的常春藤联盟。“自己的人”和“外面的人”这种标签带着丑陋的歧视，本身就是有局限的。在缺乏其他信息的情况下，贴标签可能是一种降低风险的策略，但它也排除了一个更大的潜在利益关系世界。

从某种意义上讲，排他的标签就像一个封闭的计算机环境，只有相同品牌的计算机能够相互对话，例如IBM和IBM、戴尔和戴尔。拥有沟通和行动的标准协议是有好处的，但要以可伸缩性为代价。社会技术的一项重大突破是使陌生人能够合作的开放协议：法治。法律使有着不同背景、历史、种族和社会规范的陌生人，能够以较低的风险进行交易。一项重大的投资，如买房子，即使有了财产法、建筑法规和保险的保护，也可能是令人生畏的。我们根本无法想象这些交易不受监管会是什么样子。一个社会的财富与有着执行和裁决机制的成文法律间存在着很强的联系。[30]发展经济学家在努力刺激贫穷国家的经济增长时，也认为建立法治是一个重要的障碍，缺乏强大法治系统的国家不可避免地会被贴上低效、社会分裂替代品的标签。当然，法律并不能完全取代信任，如果社会信任破裂，人们过度依赖法律机构，社会就可能无法正常运作。然而，如果没有一个运作良好的法律监管系统提供合作协议，复杂、大规模的合作就不可能进行。

沟通也是促成合作行为的关键，语言的发展极大地提升了社会和经济合作的潜力，带来了一系列新的非零和收益。关于语言是何时获得发展的，相关研究者有很多争论——公认的宽泛估计范围是距今100万年前到3万年前。[31]不管怎样，语言可能是在工具出现之后才发展起来的，250万年前的证据可以证明这一点。人类经济活动中有一个前语言时期，它的复杂程度正是因为受到了语言发展的固有限制。正如语言改变了人类创造物理技术的能力一样，它也改变了社会技术的创造。人们可以想象语言基因所具有的优势：随着词汇量的增加，有益的合作型社交游戏空间呈指数级扩大。这就好比你之前只能使用呼噜声、手势和面部表情试图和某人协商一个复杂的交易，后来你可以用50个糟糕的法语词汇去尝试沟通；现在你可以用母语流利地来完成它。

一旦发展出语言，一系列的物理技术创新就能进一步提升其在合作活动中的价值。写作大约出现在5 000年前，这使人们能够更广泛地传播知识，并随着时间的推移更准确地保存知识。我们怀疑如果没有文

字，高度复杂的古埃及、古希腊或古罗马社会是否可能存在。印刷机无疑帮助了欧洲社会从部落的黑暗时代中发展出来，而如果没有可靠的邮政服务，工业革命的社会创新是不可能实现的。要是没有电话、传真和电子邮件，现代公司也就无法建立。

由人组成的计算机

我们注意到，许多关键的社会技术创新都与信息处理密切相关，并且早期的“信息处理网络”也具有计算能力。一旦社会技术发展到了许多人可以形成合作网络、拥有大量通信和存储数据方法的阶段，人类组织就有了不同的特征——他们能够进行紧急计算。[32]

由人构成的组织有处理信息和解决个人无法独立处理或解决的复杂问题的能力。例如，英国石油公司可以被看作一台计算机，用来解决如何从世界各地提取、精炼石油和天然气，然后将其输送给数百万能源用户的问题。在英国石油公司，没有人能详细地告诉你这个极其复杂的问题是如何解决的。每天流入英国石油公司的大量数据，以及需要作出的所有决定，从董事会级别的决断到北海钻塔的轮班时间表，其中包含的信息量是无法计算的。尽管我们认为首席执行官是指挥者，但即使像约翰·布朗（John Browne）这样有能力的首席执行官也只能知晓组织中的一小部分决策，而一个大型组织在任何特定时刻都会作出成千上万个决策。然而，寻找、提取、精炼和分配石油这一极其复杂的问题能以高度精细的分配方式得到解决，日复一日都是如此。

就像蚁丘或大脑一样，人类组织也表现为一种新兴智能网络。加州大学圣地亚哥分校人类学家和认知科学家埃德温·哈钦斯（Edwin Hutchins）研究了不同环境下个体与群体的解决问题的能力。他得出的结论是，组织拥有集体的、新兴的能力，而这些能力是团体内部的个人所不具备的。[33]本质上来讲，英国石油公司不仅比其任何一个员工都聪明，而且比其员工聪明程度的总和还要聪明。

人们可能会对像英国石油公司这样的全球大公司在社会中的作用提出质疑。但即使是最强硬的企业批评家也不得不承认，像英国石油公司这样一个拥有10.3万名员工、遍布全球100多个国家的组织，是人类合作的奇迹。它的员工大多没见过面，并且永远不会见面，却被一张由社会结构、规范、协议、法律以及促使他们为共同目标努力的激励机制所组成的网络捆绑在一起。如果将合作网络扩展到英国石油公司的130万

名股东、数千家供应商和其他合作伙伴公司，那么其社会结构规模会变得更加引人注目。

英国石油公司这种规模庞大、结构复杂的组织要想存在，就必须依托进化而来的社会技术创新，包含诸如金钱等社会技术。大约在2 600年前，钱首次在美索不达米亚被使用，在本质上来说，这提供了一个通用的效用转换器，它能使一个人的经济需要和欲望以同样的单位转换成其他人的需要和欲望。[34]同样，如果没有由复式会计的社会技术（最初由意大利商人在13世纪发明）构成的财务信息中枢神经系统，英国石油公司将很难运作。[35]英国议会在1825年至1862年间通过一系列法案，确立了有限责任股份制，要是没有这个制度，英国石油公司也不可能以现在的形式存在。[36]尽管英国石油公司要依靠社会技术的馈赠来发挥作用，但它同时也是经济进化体系的参与者，它的管理者们对新的组织管理方法进行了演绎-探寻，成功的社会技术在英国石油公司之内和之外被广泛采用并传播。

我们从灵长类祖先的生物遗产中继承了为共同利益合作的倾向，在等级制度中竞争的冲动，最终获得了语言能力。随着人类尝试用各种方法组织社会经济活动，这些微小的起源在万年后发展出了演绎-探寻的进化历程。合作中固有的非零和收益奖励了那些有效的社会技术，随着时间的推移，人类发现了能使组织获得成功的技巧。在大致相关的社会技术设计空间景观中进化时，人们能够在有效技巧的基础上，让获得成功的创新在未来产生更多的可能性。创新提高了组织处理信息和解决问题的能力，开发了越来越丰富的社会技术适合度景观。同时，由于每个领域的发现都能为其他领域提供可能性，所以社会技术空间也能与物理技术空间一同进化。

我们从手斧发展到了宇宙飞船，又从狩猎队发展到了跨国公司，但人类从亚诺玛米人到纽约人的遥远路程仍没有走完，在下一章中，我们将把经济进化的最后几部分组合到一起。

13

经济是三种力量共同进化的结果

前文中概述了进化的通用模型，并将该模型映射到了经济系统的底层。在第10章中，我介绍了这样一个观点：可以将经济发展看作在所有可能的商业计划图书馆或斯密图书馆中寻找适合的设计。在第11章和第12章中，我描述了物理技术和社会技术的进化如何为经济进化构建关键的模块。在这一章中，我们将看到，随着经济进化进程在商业和市场中展开，所有这些是如何结合在一起的。

如果用进化的通用模型简化清单，我们可以看到大多数元素都有一个完整的模型。我们已经为经济进化定义了一个设计空间（斯密图书馆），一个为那些设计编码的概要（商业计划），在这些设计之下的一系列构建模块（物理技术和社会技术），一个将商业计划转换为现实的概要阅读器，以及允许让进化竞争发生的环境（市场）。然而，还有两处空白等着被填充。

首先，我们需要定义经济进化中的交互者。在生物进化中，单个的有机体扮演着交互者的角色。他们相互影响，并与环境发生作用，完成生与死的进化工作。那么，经济底层的“生与死”是什么？

其次，我们需要确认经济进化中的选择单位。在生物进化中，选择发生在基因层面。[1]那么，选择在经济背景下是如何运作的呢？

在进化经济学中，有两个问题一直以来都备受争议。[2]我将展现一个观点，它与我们遵循的进化计算观点一致，并且能融入复杂经济学的总体框架。我们将利用完整的模型深入地研究市场的运作，从而以一个新的视角看待1750年的经济创新寒武纪爆发，然后在下一章中深入挖掘财富本身的起源。

企业的“生与死”

在通用模型中，概要是交互者的创立编码。如前所述，在生物学背景下，有机体的DNA编码概要起着交互者的作用。按照我定义事物的方

式，如果商业计划是经济系统的概要，商业计划自然就能为企业的创立编码，那么经济中的交互者就一定是企业。这就引出了企业到底是什么的问题。我们也要区分企业和公司，企业是经济概念，而公司是法律实体，是一种社会技术。

在物理技术和社会技术的定义之上，我们可以用以下方式定义企业：

企业是为了营利而将物质、能量和信息从一种状态转换为另一种状态的个人或团体组织。

企业就像是将石头做成手斧，并将它们换成肉的原始人哈利，其目标是从肉中获得比生产手斧所消耗的卡路里更多的卡路里。这一定义也适用于IBM公司，为了通过出售获利，它将金属、塑料、硅、电以及工人的能量转换成计算机。这个定义也适用于服务业务。例如，擦鞋摊主可能会使用鞋蜡和通过体力劳动把脏了的鞋变得光亮，以实现其盈利目标。重要的是，该定义排除了学校烤饼义卖这样的活动。虽然烘焙销售涉及将鸡蛋、糖、面粉以及父母和孩子的能量转换为饼干，但它的目的是慈善，而非营利。

根据这个定义，企业和公司之间的区别最容易在通用电气这样的多业务公司中体现出来。通用电气公司是一个单一的企业实体，但在塑料、照明、媒体、金融服务等领域有着独特的业务。大多数大公司实际上都是企业的集合体，例如可口可乐公司根据地理位置和饮料产品来划分分公司。而许多较小的公司只是提供单一产品或服务的企业，例如“十八街新闻杂志”可能只拥有一家街角小店。

我们可以这样定义公司：公司是由一个人或一群人共同控制的一个或多个企业。

现代公司呈现出不同的法律形式：独资企业、合资集团和企业集团。但即使没有这些正式的社会技术，我们也可以说哈利是他的“手斧制造企业”的唯一所有人，因为他实际控制着这家“企业”。同样，亚诺玛米人的狩猎队可以看作一个控制打猎业务的合伙企业（请注意，我使用了“控制”一词，而不是“拥有”，因为所有权意味着财产权的社会技术）。我们可以说通用电气是一家公司，因为它所有的多元化业务最终由同一个股东集团控制。

考虑到前面的定义，我认为企业是进化系统中的交互者，而不是公司。顾名思义，交互者是进化系统中的实体，它们相互作用，并与环境

相互影响，在环境的选择压力下经历不同的成功率。对于业务单一的公司来说，企业和公司的区别并不重要。企业和公司本质上是相同的：如果“十八街新闻杂志”街角商店这一企业成功或者失败了，那么“十八街新闻杂志”街角商店这一公司也就成功或者失败了。但在通用电气公司的例子中，与客户、供应商和竞争对手的互动主要发生在各个业务部门的层面，而不是公司层面。在塑料行业中成功或失败的不是通用电气公司，而是通用电气塑料集团。如果通用电气塑料集团的业务失败了，它会对通用电气公司的财产造成损失，但通用电气作为一个公司很可能会生存下来，因为它能继续从事许多其他业务。

下一个层次是要区分企业产品服务和个人产品服务。例如，有的人可能会争论通用电气塑料集团到底是一个企业，还是一个包含聚碳酸酯和改性聚苯醚等多项业务的实体。一个企业的关键属性是它能为交互提供焦点，因此，要区分企业和产品服务线，我们应该寻找不同维度中的一致性，比如说共同的客户、竞争对手、地理位置、技术或供应商。街角商店是一个单一的企业，而不是牛奶企业加上报纸企业，因为相同的客户倾向于同时购买牛奶和报纸，该店的主要竞争对手是其他街角商店，而互动的焦点是在商店层面。同时，一家制药公司可能拥有心脏药物企业和癌症药物企业，其中每个企业都有独特的客户、竞争对手和技术。

划定这些界限时不可避免地会存在一些模糊性和主观性，而业务部门和产品或服务线的构成是许多公司经常讨论的主题。我们将采用实际的方法，假设最好用管理团队对它们进行区分。因此，我们会说，经济进化中的交互者是由公司定义的业务部门（我们将由通用电气公司决定聚碳酸酯是业务还是产品）。这种方法的优点在于可以捕捉到边界的动态性。随着环境的变化以及技术、客户和竞争对手的转变，管理团队会定期重新评估和更改业务部门的边界。

选择单位

企业是经济进化中为了生存而努力的交互者，但这并不意味着企业是进化选择的单位，就像我们不能假设生物系统中的选择发生在有机体上。[3]为了充分了解进化是如何进行的，我们需要再降低一个层面。

这里我们需要回顾一下之前有关进化逻辑的讨论，也就是“好的复制因子能进行复制”。在进化的通用模型中，选择单位是为“性状”编码的概

要，这些性状能使一个交互者与其他的交互者区分开来，从而在竞争中留存并复制。在生物的例子中，这些小的概要是基因。例如，如果一组特定基因能比其他基因给有机体创造更好的伪装，而这种伪装能让基因的携带者在生存和复制中具有独特的优势，随着时间的推移，这组基因在群体中出现的频率就会增加。

经济进化中的选择单位是什么呢？考虑到选择单位都是在选择概要，所以它们一定是商业计划中的一小部分。但是是哪部分呢？我们怎么识别它们呢？它们又位于什么粒度级别呢？选择单位是营销策略，还是特定产品的物理技术呢？

要回答这些问题，我们就要及时回顾，并且探究让企业的竞争成功与众不同的特质是什么？如果商店形式的差异能在一段时间内使博德斯书店的市场份额比巴诺书店的市场份额多一些，那么在这期间商店形式就是选择单位。如果库存管理系统的差异不会带来影响，那么它就不是一个选择单位（尽管它可能在未来某个时候成为一个选择单位）。

如果说选择单位就是环境选择的东西，听起来会有种循环的感觉。但我们别无选择，适合度函数是复杂、多维且随时间变化而变化的。我们不能先验地说出系统正在选择的是什么，只能回顾性地观察选择，因此只能采用经验性的、回顾性的方法去定义选择单位。

这种回顾性的、经验性的方法在生物学中同样重要。基因的定义也具有模糊性。人们普遍认为，决定长腿等性状的基因是专门负责编码的某种DNA序列。然而，现实情况是，没有专门的、有序的DNA序列只编码长腿这一特征。DNA不像一个简单的蓝图，它更像一个密集而复杂的交换网络。长腿性状是由许多相互作用的子模块组成的，这些子模块都具有潜在的遗传性，例如调节骨骼生长的激素和肌肉构造过程等。这些模块也由子模块组成，如此层层细分，直到达到单个蛋白质的层面。这些相同的模块也可能参与到其他特性的构建过程中，因此，生物学家在基因鉴定中采取了一种实用和经验性的方法。例如，一些疾病是由某种蛋白质缺失引起的。科学家们会寻找负责编码该蛋白质的DNA序列，然后说他们已经找到了该疾病的基因，但实际上特定的DNA链也可能参与到许多其他机制中。基因仅仅是一个方便的标记手段，用于标记散布在DNA上的一些概要片段，并参与遗传性状的编码，这些性状可以为生物的生存和繁殖提供基础。

我们将采取类似的实用和经验性的方法来确定经济发展中的选择单位，并定义一个新的术语——模块。模块是过去提供的或未来可能提供

的商业计划的一个组成部分，它是竞争环境中对不同企业进行差异化选择的基础。

思考模块的一个方法是假设："如果我是一个企业的经理，我可以尝试改变什么来提高企业的绩效？"你可能会想到发起新的销售活动，重新设计客户服务流程，引入新的成本控制规则，或者增强产品研发。为区分业务绩效提供基础的任何活动都是模块，或者更准确地说，为这些活动编码的商业计划组成部分都是模块。

有强有力的证据表明，人们能够回顾性地识别模块。整个顾问行业、商学院教授以及确定和传播商业"最佳实践"的专家，实际上是4个模块（"最差实践"也是模块，只不过是不适合的模块）。[4]这些研究人员研究公司，以便寻找过去带来成功的商业实践，并在书面文件和口头陈述中描述这些实践（例如将它们编码成概要），将它们传递给管理团队（例如合格的概要阅读器）。管理团队会尝试在其组织中实施这些实践，以提供不同的竞争优势来源（例如作为一个选择单位）。

像其他定义一样，这个定义也带有不可避免的主观性。企业之间的绩效差异往往是由许多复杂的原因导致的，就像有机体对适合度的贡献是极其复杂的一样。正如我们将在第15章中讨论的，过去最佳的选择单位实践模块可能不是未来优势的来源。进化选择并不是整体上作用于商业计划，而是作用于商业计划中的元素的。

选择单位的问题是经济学和生物学界争论和研究的主题。多年来，人们对经济系统中的选择单位，例如理查德·纳尔逊和悉尼·温特的"惯例"，以及更普遍的社会系统，如理查德·道金斯的"模因"或罗伯特·博伊德（Robert Boyd）和彼得·理查森（Peter Richerson）的"文化变体"，提出了各种建议。[5]每一个术语都有其各自的优势和弱点，而模块的定义借鉴了这些概念。我引入一个新术语是为了确保它与我们正在建立的进化通用计算模型一致，并澄清经济环境中的模块、交互者和选择单位之间的区别。

策略胶水

"模块"一词还有着更深的内涵。正如之前讨论的，进化概要有一个组建块，使得概要具有组合的特征，模块一词意味着商业计划是模块的集合，人们可以通过组合并匹配各种模块来创造不同的商业计划。如果人们打开一份真正的企业商业计划，看一看目录，就会发现标题中有着

诸如此类的词语。[6]

- 市场环境。
- 策略。
- 产品和服务。
- 运营。
- 市场和销售。
- 组织。

这些商业计划都包含子组成和子子组成。例如，化工企业的商业计划中可能会有一个产品模块描述企业将如何为航空行业提供特殊的碳纤维材料。该模块将建立在一系列用于制造碳纤维材料的物理技术之上。同样，市场营销和销售可能会有一个直接让销售人员销售这些产品的模块。该模块建立在许多社会技术之上，用于组织、管理和激励销售人员。

商业计划是从原子级的物理技术和社会技术上建立起来的，这些技术结合成了模块，模块又结合成了商业计划，使它们结合在一起的胶水是策略。我们将在第15章中更集中地讨论策略，但是现在，我们可以将策略视为一个假设，即在给定的环境中，什么样的模块组合是有利的。例如，一位企业家可能会假设，从苹果公司借来电子设备这一物理技术，再从戴尔公司借来网络销售和物流的社会技术，然后将两者结合起来，用一种适合的设计创造出来的新玩意儿就会是一个成功的组合，这样的假设就是策略。这个企业家之后可以写一份描述这一策略的商业计划书，然后再冒险向硅谷的资本公司寻求资金。商业计划的主要内容是在特定的策略之下，使物理技术与社会技术结合起来，成为经济世界中真正的企业。

变异：从企业家到官僚

进化游戏中的玩家现在已经就位，是时候加入情节，看看进化过程是如何在经济基础上展开的了。在前面的几章中，我们看到了演绎-探寻是如何提供机制，区分物理技术和社会技术空间的。同样的原则也可以用在商业计划的区分上。经理们尽己所能，提出了他们希望能够成功的

计划。但是，詹姆斯·柯林斯（James Collins）和杰里·波拉斯（Jerry Porras）在研究长期存在的公司的《基业长青》（*Built to Last*）一书中提到，要“尝试大量的事情，然后保留有效的”。[7]在任何时候，商业计划的实验都有着惊人的多样性，从沃达丰在英国发明的3G移动互联网服务，到英国石油公司在俄罗斯探索其业务选择，再到玻利维亚一家街角小店更换了一种新的豆子品牌，看顾客是否喜欢。商业计划空间中的演绎-探寻为选择行为提供了巨大的差异性或超凡的繁殖力。像开发新的社会技术一样，开发成功的商业计划更像是艺术而非科学。虽然它不像生物进化那样完全是一个盲目的过程，但商业计划差异化仍然需要更多的探寻，而不是演绎。

这个范围内的其中一端是那些创造全新商业计划的企业家。创业过程就是一种演绎-探寻的形式，因为创业者会采用各种商业计划的模块，以新的方式组合它们，或将新的物理技术或社会技术引入商业世界。例如，家得宝（Home Depot）的创始人采用了郊区超市的零售模式，并将其与“自己动手”的建筑用品相结合，制订出一个新颖而成功的商业计划。

伟大的英国经济学家阿尔弗雷德·马歇尔把企业家比作中世纪的骑士，并称赞他们是资本主义的英雄。[8]但是，从进化的角度来看，企业家并不是经济创新的唯一来源，因此也不是唯一的英雄。中层管理者也可以成为英雄。想象一下，一位来自兆丰有限公司12号楼23号房间的中层经理，刚刚提议了一个新的程序来管理产品保修索赔。这位中层管理者采用了大胆的企业家所使用的演绎-探寻程序，他试图理性地计划自己的行动，然后从环境（包括客户和老板）中得到反馈，并进行调整和尝试。从进化的角度来看，这位中层管理者和企业家一样重要，也是商业计划多样性的来源，甚至可能不止于此。

适合度景观的阿尔卑斯山形状意味着进化需要一系列的跳跃来实现有效搜索。大多数创新都应该是小而短的跳跃，因为绝大多数激进的跳跃创新都是失败的。23号房间的中层管理者的渐进式创新比企业家的大胆计划更有可能获得成功。但我们也需要一些中长距离的跳跃来防止进化卡在局部的山峰上。因此，进化既需要呆板的官僚，也需要狂热的企业家。

选择：大男人对市场

企业家、经理和官员通过演绎-探寻实现了商业计划的差异化，那这些计划是如何被选择的呢？在历史上，人类发展了两种经济选择方法：大男人和市场，[9]我们将依次讨论这两者。

在经济发展的早期，选择过程相当直接——生存。如果你选择了一个策略，如在河边狩猎黑斑羚，并结合了物理技术（如弓和箭）和社会技术（如狩猎队），获得了成功，那么你的卡路里收入就会大于支出。这种卡路里利润能够让你做一些事情，比如为儿童投资。通过吸引更多的参与者，在卡路里上盈利的商业计划将拥有更高的被复制机会，并被下一代所采用。其他的商业计划，例如在平原上使用弹弓猎杀鸵鸟，如果它们的卡路里利润没有那么高，那么它们将失去通往更成功计划的资源，并随着追随者转向不同的计划或追随者本身的消亡而消亡。

随着社会和经济变得越来越复杂，选择的反馈回路变得越来越不直接，中间的、由社会驱动的选择出现了。毫无疑问，当第一个大男人说："让我们把这片肥沃的土地给我第三任妻子的表弟（一个糟糕的农民），而不是给X先生（一个优秀的农民）"，选择和社会之间就会出现冲突。我们可以放心地假设，政治干预经济事务与政治和经济本身一样古老。这种倾向于糟糕的商业计划（表弟）而不是好的商业计划（X先生）的决定不会持续太久，在这样一个环境中，如果大男人经常这么做，那么其所属的部落要么在他的领导下灭亡，要么大男人在叛乱中被推翻。但是，一旦一个社会跨过了生存门槛（特别是在定居农业出现之后），这种选择过程中的社会短路就不仅是可能的，并且随着该群体的日益富裕，它的可能性会变得更大。

如果一个部落生存了下来，而且大男人的贪污、腐败或能力并不会威胁其发展，那么只有相对较少的人能意识到他们部落所放弃的额外财富。正如罗伯特·赖特指出的那样，竞争对这一点有一些制衡。最终，另一个大男人可能会承诺做得更好，然后推翻旧的大男人，或者表现不佳的部落可能会被组织得更好的部落通过暴力接管。[10]但是，没有什么可以保证新的大男人或部落会比旧的好。在商业计划的选择过程中，政治干预的主要影响是减缓进化的进程。在极端情况下，首领、国王、独裁者和其他大男人实际上可以阻碍经济发展的轨迹，只要人们仅仅是接近饿死，而不会真正饿死，这种进化的死胡同就可能持续很长时间。

商业计划选择的"大男人"制度还有一个问题：大男人扭曲了适合度这一功能本身。进化算法的一个特点是，它们能够很好地适应给定的任何适合度标准。正如我们之前看到的，当计算机研究人员卡尔·西姆斯选

择适合游泳的人工生物时，他惊讶地得到了一群聪明的游泳运动员。当他把适合度标准改为在陆地上穿行时，研究对象的鱼鳍和尾巴都消失了，取而代之的是腿和蛇形的身体。当工程师们借助人工进化来设计半导体、软件或新药时，他们必须非常小心地指定适合功能，因为错误的适合功能必然导致错误的设计。在大男人制度中，适合度函数最大化地体现在大男人及其亲信的财富和权力上，而不是社会的整体经济财富。群体中创造性的、企业家的和演绎-探寻的力量都指向取悦大男人，从法国的城堡到俄罗斯的冬宫，遍布世界各地的无数豪宅和宫殿都在以其奢华的展示取悦游客，这证明经济发展在最大限度地发挥大男人财富的适合度函数方面具有有效性。

到目前为止，人类唯一想出的替代大男人的选择系统就是市场。市场是一个既古老又现代的发明。一方面，自由市场早在哈利和拉里第一次交换斧头和肉时就已经存在了。尽管这种非正式的双边贸易在历史上一直存在，但有证据表明，有组织的市场最早是在近代出现的，它可以追溯到公元前7 000至公元前5 000年新月沃地农业的发展以及乌尔和巴比伦等城市的崛起。[11]

传统经济学的一大成就是它表明市场试图满足的适合度函数，实际上是市场所有参与者的整体福利。在大男人经济中，企业因政治利益得以生存或消亡。在以市场为基础的经济中，企业的生死取决于客户是否喜欢并愿意为其产品和服务付费。在大男人经济中，资源被引向了那些符合大男人利益的企业。在市场经济中，资源被引向了那些能充分利用资源产生经济效益的企业。

历史上所有的经济体事实上都是大男人和市场的结合。过去生活在狩猎采集社会中的人能够自由交易，只要他们的交易不会冒犯大男人，并保证大男人能得到他的份额。巴比伦和古雅典安哥拉市场上的商人通过政治联系、税收、贿赂和其他等级制度来获得交易许可。

人类历史的大部分时期都是大男人主导着商业计划选择，而市场则扮演着次要或秘密角色。例如，经济学家威廉·鲍莫尔（William Baumol）估计，欧洲的封建贵族阶层控制着80%以上的财富总量。[12]直到大约300年前，大男人和市场之间的权力平衡才开始转变。我们接下来看看选择在基于市场的系统中是如何运作的。

选择如何在市场系统中运作

如果说选择在大男人经济中是相对简单的，它在市场经济中就相对复杂一些。从进化的角度来看，定义市场经济的特征是，市场是挑选商业计划的最终力量。这并不意味着市场经济中没有大男人的等级制度，相反，从马克思到现代反全球化抗议者，任何人都能很快地指出，资本主义社会中充斥着强盗式贵族、公司总裁和其他有钱有势的人。大型现代企业的组织结构图与其他任何大男人阶层的组织结构图几乎没什么不同，尤其是大多数企业仍是由大男人主导的，大女人并不是很多。然而，在资本主义社会中，一个人变得有钱有势的方法是拥有市场青睐的商业计划，而不是其他。

基于市场的经济体运行着双层的商业计划选择系统。大多数的经济决定仍然是在等级制度下作出的。正如商业历史学家艾尔弗雷德·钱德勒（Alfred Chandler）在20世纪70年代观察到的那样，企业等级制度这只“看得见的手”比亚当·斯密提出的市场这只“看不见的手”作出的经济决策要多得多。[13]但是，在这些庞大的等级制度之上的是一个单薄但关键的层面，在这个层面上，等级制度会与市场相遇。市场经济是进化的竞争等级制度。

前文中已经讨论过演绎-探寻是如何区分不同商业计划的，接下来让我们更深入地了解一下差异在市场环境中的运作方式，以及它与环境的交互如何决定商业计划的选择。

假设有一位主管负责一家大公司的一个业务部门，她正在考虑修改自己的商业计划，从而提升部门的利润。她所做的第一件事就是为企业创造各种选择：“我可以尝试扩大产品线A，引进新的服务B，在这里削减成本，在那里进行重组。”每个选项都涉及商业计划中不同的模块集。在生成这些选项的过程中，她运用了所有的认知能力：一些想法来自对业务的演绎思考；一些来自分析以及从其他经历中获取的经验，例如“这在我以前的公司中有效”；一些想法来自模仿，例如“我的竞争对手推出了这个产品，并且成功了”；还有一些想法是他人传给她的，例如“市场营销部的戴维提到了这一点，可能是个好主意”。

一旦这个业务部门主管有了一组选项，她就会进行一些心理模拟，以思考哪些选项是最好的。在头脑中筛选出选项后，她可能会进入下一个选项生成和测试阶段。她可能会要求同事提供更多的想法，或者要求他们对她的想法进行心理模拟，然后再给她反馈。接下来，这位主管和她的同事们将创建一个更精确的选项列表，并进行更正式的模拟。例如，他们可能会创建电子表格模型，尝试不同方案，分析成本，或委托

其他公司或部门进行市场研究。团队将再次评估这些选项，或许母公司的高层管理人员也会讨论这些选择。最终，业务部门主管将选定一个或多个选项，作出决策，并将现有商业计划更改为新的计划。这将导向现实世界中的行动：分配预算和人员，改良产品，调整销售活动等。一旦这些行动在现实世界中实现，市场就会作出判断，销售业绩会上升或下降，利润会增加或减少，因此管理团队会收到有关其商业计划成败的反馈。

这个过程并不总是如此正式或如此清楚的，但基本包括选项生成、测试和选择的迭代循环。这个循环从单个行为者的心理模型开始，最终在一组行为者中传播开来，然后导向现实世界中的行动。选择会在多个层次上起作用：在单个行为者的心理模型中、在组织等级中以及在市场中。

复制：放大成功

进化算法的最后一步是复制。在生物学进化中，复制通过细胞分裂或有性生殖进行。对生物体生存和繁殖能力有益的基因往往会被复制。因此，复制的效果是，随着时间的推移，适合的基因在人群中出现的频率增加了，不适合的基因频率降低了。在生物系统中，基因频率是通过人口拥有特定基因的比例来测量的。通过测量频率，我们就可以知道在交互者群体中有多少交互者拥有某个特定的选择单位。

如果我们将同样的成功复制测量方法应用到经济系统中，就可以了解到企业中包含某个特定模块的百分比是多少。我们在测量中会遇到一个问题：在大多数物种中，交互者的大小变化不大。例如，成年人类的身高在1～2.4米之间，变化范围小于一个数量级。然而，经济领域的互动者——企业，在规模上的变化可能是从一个人维持的擦鞋摊到埃克森美孚的上游石油业务，差距大约是109个数量级。由于某种原因，我们需要说明一个事实，即在5个擦鞋摊复制的一个商业计划模块与在5家主要石油公司复制的一个模块具有不同的经济效应。

在生物学中，频率的概念实际上是一个更基本的度量方式的抽象。与其问拥有某个基因的生物比例是多少，不如问问物种总生物量中含有该基因的比例是多少。这种方法更符合以基因为中心的观点，因为它问的是有限的化学和能源资源（以生物量衡量）中有多少是被特定基因控制或影响的。对大多数物种来说，每个生物体的平均生物量相差不大，

我们可以将群体的频率或百分比视为一个更容易测量的替代物。当我们问到人类生物量中受蓝眼睛基因影响的适合度百分比时，答案是约20%。这种资源百分比的观点也更符合我们用中立的基础研究进化的方法，因为它可以应用于从计算机到乐高积木的各种系统。

我们将采用资源百分比的方法，并且通过特定商业计划模块影响其适用性的资源百分比来衡量复制的成功。这些资源可能包括金钱、劳动力、厂房和设备，甚至像品牌、技术知识和客户关系这样的无形资产，即任何有限供应和企业竞争的资源。[14]既然现在讨论的是连续的而不是离散的度量，那么我们会说成功的商业计划模块被放大了，而不是被复制了。如果商业计划模块对资源的影响随着时间的推移而增加，那么它将在业务计划空间中得到放大。[15]

现在，我们可以回顾一下有关选择的讨论，看看成功的商业计划模块是如何被放大的，以及如何受到奖赏进而影响更多资源的。回到假想的业务部门主管的案例中，当主管为自己的商业计划考虑各种选择时，她做了决定，并选择了她认为会成功的模块的特定配置。主管通过实施计划并配置人员、资金和其他资源来奖励这种配置，因此，她选择的特定模块得到了放大，对资源有了更多的控制。主管考虑和放弃的其他10个选项没有得到奖励，至少在现有的业务中，这些模块已经不复存在了。我们假设她的商业计划的早期结果看起来非常有希望，因为销售额增长了，利润也上升了。公司高层团队对她印象良好，希望在公司的其他业务中尝试一些相同的模块，例如新的销售方法。更多的资金和人员被投入这些模块，因此这些模块被进一步放大。

几个月后，这个主管的一个竞争对手注意到这些变化并复制了这些模块，从而使它们进一步扩大，现在这些模块也影响到了第二家公司的资源。在一定程度上说来，这个模块的成功使两家公司都有所扩大，从银行和股票市场吸引到了更多的资本，模块的影响再次被放大。后来，这个业务部门主管决定购买一家尚未采用该模块的小型公司，因为这是实施该模块并使目标公司运营最佳的好机会。交易完成后，该模块在被收购公司实施，再次扩大了其影响范围。回顾生物学准则“好的复制器能进行复制”，人们可用“模块之眼”观察经济世界，从而看到“好的放大器能放大”。

我们可以看到各种用于传输和扩展商业计划模块的机制，它们包含个人的心理模拟和决策，团队的问题解决和决策，组织内外的模仿，市场对成功模块的资源配置，以及一个企业接管另一个企业。

简述经济进化

我们现在有了在斯密图书馆中寻找适合的商业计划进化模型中的全部元素，也已经看到了商业计划是如何指导创建企业的，这些企业都是可由合格的商业计划阅读器实现的，这些指导将物理技术和社会技术结合到了同一策略下的模块中。行为者在寻求可能有益的计划时会通过演绎-探寻法来区分不同的商业计划。虽然这个过程的实验分布不同于纯随机的生物进化分化，但它仍然为进化算法提供了具有超凡传播力的商业计划，使选择得以发生。

选择过程是嵌套的，并且发生在多个层次，从个体的心理模拟到群体解决问题的活动。当商业计划向上和向下渗透到组织的层级中时，就会发生进一步的选择，然后在某个点上，计划会被实施，市场会作出判断。

成功的模块可以获得对更多资源产生影响的回报。模块的成功体现在两个层面。第一个层面在组织内部，尤其是商业计划得以实施并获得资源时，例如，投入人力和财力来实施一项计划。第二个层面是模块在市场中表现出来，通过客户和金融市场获得增长和更多资本。这是一个筛选的过程，商业计划的超凡传播力意味着一个组织考虑的计划远多于可以实施的计划，并且在市场中进行实验的计划也远多于能成功的计划。随着选择的展开，适合的模块最终会对更多的经济总资源基础产生影响。这是一个高度动态的过程，今天适应的，明天可能适合也可能不适合。随着模块的更迭和企业的兴衰，进化过程永不停止，不断地适应着市场的需求。

在进化算法的基本机制之上，我们不会期待经济进化和生物进化有任何特别的相似之处。[16]例如，在生物进化中，选择单位这一概念在分散的代际间遵循着“随着修改而下降”的原则，但在经济进化中，从一个商业计划跳到另一个商业计划并不会使进化过程变得更为缓慢，只是会有所不同。同样，人类使用大脑并拥有远见意味着分化和选择的机制在经济系统和生物系统中是非常不同的，但它们都是进化而来的。

此外，在我提出的框架中，没有任何东西是以有组织的市场、货币、私人财产、公司，甚至写作等现代社会技术的存在为前提的。同样的过程也适用于原始人哈利，他思考如何用手斧换取更多的肉，就像现代跨国公司的高级总经理考虑在中国实施何种战略那样。

赞扬市场——以不同的理由

如果说有一件事是传统经济学家所认同的，那就是市场是好的——市场可能并不总是完美的，确实有一些很明显的缺陷，但是当它们发挥作用的时候，又是很难被击败的。这一建立在一般均衡理论基础上的结论是20世纪“冷战时期”意识形态争论的中心，在今天仍在为全球资本主义提供知识基础。经济进化的观点使人们对市场是好的这一观点达成了共识，但背后的原因大相径庭。

传统经济学强调，在均衡条件下，市场是以优化社会福利的方式分配资源的最佳方法。[17]正如我们所见，问题在于，现实世界从未达到均衡。如果传统经济学的均衡假设是错误的，我们怎么能确定市场是好的呢？

按照我刚刚概述的框架，我们可以将市场解释为一种进化的搜索机制。市场为差异化的演绎-探寻过程提供了激励。关键是，市场之后提供了适合度函数以及能代表人类广泛需求的挑选过程。市场提供了一种将资源从适合的模块中转入，从不适合的模块中转出的方法，从而扩大了适合模块的影响。

简言之，市场运行良好的原因可以归结为进化论理论家所说的“奥格尔第二规则”（以生物化学家莱斯利·奥格尔［Leslie Orgel］命名），即“进化比你更聪明”。即使是一个高度理性、聪明、仁慈的大男人，也无法在生态系统中寻找高峰这方面打败进化算法。市场在指挥和控制方面更具优势，并不是因为它们在均衡中配置资源的效率，而是因为它们在非均衡中创新的效率。

复杂经济学家不会忽视市场的分配作用，[18]他们也倾向于认为市场在这方面相当擅长，而且通常比大男人做得要好。传统经济学家喜欢谈论均衡市场的“完美效率”，复杂经济学家往往将市场效率视为一个相对的概念。复杂经济学理论认为，拥有完美效率的理想状态可能并不存在，即使存在，市场的不均衡性也可能阻止它的实现。市场永远无法实现完美的效率，就像汽车的引擎永远无法达到100%的热动力利用效率一样。市场善于分配更多和它们作为分布式处理系统的计算效率有关，例如将正确的信号发送给对的人，而与它们达到理论的全球均衡的能力无关。[19]

市场的经验记录表明，其成功是一种进化机制。正如经济学家威廉·

鲍莫尔所说，自由市场历来都是“创新机器”。[20]你周围有多少个库存单位是在大男人经济体中发明或设计的？几乎没有例外，现代世界大量的物理技术和社会技术创新都来自市场经济。

这些并不是说市场导向的社会是完美的。富裕资本主义世界的社会存在着严重的问题，比如不平等、环境破坏和艾滋病等危机。有强有力的证据表明，这些社会中盛行的唯物主义并不一定能让人们更快乐。[21]贫穷的大男人社会也面临同样的问题，但通常情况会更严重，因为可用于解决这些问题的资源较少，而且新的创新方法解决问题的可能性较低。市场的进化观并没有使许多国家在从大男人经济向市场经济转型过程中所遇到的困难变少，而这些经济转型不可避免地会涉及扭曲的社会变革。在第18章中，我们将看到复杂经济学也为审视市场局限性和弱点提供了不同的视角。人们往往会用脚进行投票，特别是在现代，世界范围内的移民都是从大男人经济体到市场经济体的。

市场的复杂性促使人们欣赏推动创新发展的市场优势。也许对大男人经济体系最有说服力的批评之一来自瓦茨拉夫·哈维尔（Vaclav Havel），他在监狱里度过了一段时间，后来成了捷克民主共和国的总统，哈维尔说：“生活的本质是无限且神秘的，它的完整性和可变性无法被中央头脑遏制或规划。”[22]

元创新：回顾1750年

在这本书的开头，我提到过人类历史上最令人吃惊的事件是开始于1750年的财富和经济复杂性大爆发，这一爆发一直持续到现在。通过进化框架，我们不仅可以看到发生了什么，还可以看到它是如何发生的以及为什么发生。在这一时期，社会技术的一系列创新显著提高了经济进化本身的速度。

在第11章中，我们讨论了第一个元创新，即科学革命。在1500年以前，人类的知识主要是通过试验和犯错建立起来的。科学的出现加快了通过物理技术空间进行搜索的步伐，极大地提高了理性演绎和实验探寻的有效性。[23]

第二个元创新是有组织的市场的兴起。市场经济的发展并非发生在大爆炸中，而是两个世纪以来社会技术进化的结果。引发这一变化的两个关键事件是英国议会民主的建立和美国革命。

在1509年到1547年之间，英国的亨利八世是其王国的绝对统治者。但是到了1690年，经过一个世纪的动荡和改革，英国成了一个君主立宪制国家。议会控制了政府的钱袋，并成立了英格兰银行来管理国家货币。[24]此外，英国还采取了重要措施来保障法治，保护个人财产。1628年的《权利请愿书》（*The Petition of Right*）和1679年的《人身保护法》（*Habeas Corpus Act*）尤为重要。所有这些基础的社会技术变革使经济在18世纪中从封建、等级森严的大男人社会转变为市场经济，商业阶层不断壮大，私营企业竞争激烈，资本市场不断扩大。尽管其他欧洲大国在它们的经济中保留了更多自上而下的控制，但机构重组也在一些其他国家中发生，尤其是北欧国家，这使得市场力量在整个欧洲大陆缓慢扎根。

1776年开始的美国革命是市场经济发展的一大推动力。历史学家保罗·约翰逊（Paul Johnson）指出，英国人在经济上统治殖民地的方法与其他欧洲大国大不相同。[25]法国、西班牙和葡萄牙都将高度集中的、等级森严的政治、军事和宗教指挥体系移到了新的殖民地，这又导致了新世界经济体中自上而下的控制。而英国人十分吝啬，不愿投资构成这种控制的军事力量。大英帝国其他地方存在着更为紧迫的忧虑，比如英国与欧洲大陆的关系一直很复杂，因此，英国在美洲的殖民地演变成了高度独立的农场主和商人的分散聚居地。从那时起，英属北美殖民地就成了一个拥有自由贸易、自由思想和自由宗教的地方。每当英国或殖民当局试图镇压时，居民们只会前往任何远离控制中心的荒地。

到1776年革命来临时，美国已经经历了经济相对自由的一个世纪，并发展出了平等主义以及一种天然质疑大男人控制的平民文化。殖民地也有大量富有的中产阶级。约翰逊指出，1750年，典型的美国东海岸农民拥有大片土地、10头牛、16只羊、6头猪、2匹马和一队公牛。一般的家庭有6～7个孩子，其中通常有4～5个能活到成年，只有3%～5%的中年男性非常贫穷。[26]到了革命时期，美国的自由市场实验取得了成功：1700年，美国的国内生产总值是英国国内生产总值的5%，到1775年，美国的国内生产总值是英国的40%，正如约翰逊所说，“是世界上有史以来最高的增长率之一”。[27]

到18世纪末，社会技术已经为北美、英国以及北欧其他地区的市场经济奠定了不同程度的基础。这些仍然不是我们今天承认的市场体系，例如，那时的银行系统相当原始，有限责任股份公司还没有发明。但是大男人已被排除在商业计划选择之外，创业精神蓬勃发展。

科学也深深扎根于这些地区，因此，这些地区成为工业革命的焦点就不足为奇了。在19世纪和20世纪，新物理技术的发展速度大大加快，市场促使商业计划的进化加速，将物理技术转换成产品和服务。物理技术、社会技术和商业计划创新的良性循环跟着到来，通向了目前为止世界上最强盛的经济增长期。

14

重新定义财富

1948年，尼古拉斯·乔治斯库-罗根（Nicholas Georgescu-Roegen）和他的妻子满怀恐惧地藏在一只桶里，登上一艘从罗马尼亚出发的货船，偷偷前往伊斯坦布尔。[1]如果被抓住，他们就会被杀死。

乔治斯库-罗根是一名经济学家，一直在为第二次世界大战后的罗马尼亚政府工作。他曾在巴黎的索邦大学接受培训，在20世纪30年代回到罗马尼亚之前，曾在哈佛大学约瑟夫·熊彼特身边学习。战后的几年里，罗马尼亚十分混乱，充满了有关国家未来的权力斗争。对乔治斯库-罗根来说，离开的时候到了。

罗根和妻子最终来到了美国，在那里，他得到了在纳什维尔的范德堡大学教员的职位。在接下来的20年里，乔治斯库-罗根对传统经济学理论作出了开创性的贡献。他的数学造诣使他成为一颗冉冉升起的新星，新古典主义经济学家保罗·萨缪尔森授予他终极荣誉："经济学家中的经济学家。"但在1966年，60岁的罗根开始猛烈抨击传统理论，变得叛逆起来。[2]此后，他通过研究进化论和物理学来寻找传统经济学的弱点，并于1971年出版了巨著《熵定律与经济过程》（*The Entropy Law and the Economic Process*）。[3]

在这一章中我们将看到，乔治斯库-罗根的思想远远领先于时代。《熵定律与经济过程》一书的基本观点是，经济活动在根本上是关于秩序创建的，而进化是创造秩序的机制。我们将研究乔治斯库-罗根开创性的想法，将它们与当前的科学联系起来，并将它们与前几章概述的进化模型结合起来。这样就能到达我们的目的地：有关财富起源的新视角。

怪人和不成熟的投机者

乔治斯库-罗根认为，虽然人类物种的生物性状是通过基因缓慢进化，或者说是"在体内"进化而来的，但同时我们也通过文化快速地"在体外"进化。乔治斯库-罗根不是第一个观察到这一现象的人。达尔文将此

视为自己的理论的一个启示，20世纪50年代，天主教神学家皮埃尔·泰尔哈德·德·夏尔丹（Pierre Teilhard de Chardir）基于体内和体外进化的思想发展出一种哲学。[4]乔治斯库-罗根也不是当时唯一一位通过文化进化寻求答案的经济学家。弗里德里希·哈耶克在1960年出版的《自由秩序原理》（*The Constitution of Liberty*）一书中写到了文化进化，肯尼思·博尔丁在20世纪70年代发表了他的文化进化和经济进化理论。[5]乔治斯库-罗根在科学的基础上建立理论的能力是独一无二的，尤其是在进化论和热力学之间的联系方面。

在第3章中，我们讨论了热力学第二定律，即宇宙会不可避免地从低熵状态转变为高熵状态原理。如果放任不管，世界就会从有序走向无序。但是，如果你把能量投入一个开放系统，就可以暂时对抗熵的上升，在宇宙的局部创造秩序。然而，第二定律必须发挥它应有的作用，任何熵值下降的开放系统最终必须以热量和废物的形式将它的熵输回宇宙，从而使宇宙的总熵继续上升。这也意味着一个开放的系统必须有持续的能量输入，从而不断地对抗熵；如果能量输入被切断，秩序就无法保持，系统就会衰退并消散。

热力学第二定律从根本上塑造了生物世界的进化。[6]从热力学的角度来看，生物体是高度有序的分子的集合。所有生物都具有某种形式的屏障，如薄膜、皮肤或外壳，从而使其有序的内部与无序的外部隔离开来。保持这种内外的区别需要能量。我们可以说，边界内的分子是有序的，因为仅从随机分子运动中获得生物体内特定化学模式的概率是无穷小的。例如，随机移动的分子极不可能自发地组成一种功能齐全的细菌。纽约大学的生物化学家罗伯特·夏皮罗（Robert Shapiro）将此比作垃圾场的龙卷风组建一架波音747的概率。[7]因此，每个生物体都需要一个能量源来维持和发展其复杂的内部秩序，当熵被回馈给宇宙时，所有生命都会释放热量和废料。当这个过程停止时，生物体的分子又回到了无序状态——死亡是对第二定律的屈服。

热力学第二定律为所有生命提供了一个基本限制：随着时间的推移，能量输入必须大于能量消耗。所有的生物体都必须通过热力学“盈利”来生存和繁殖。生物体的设计可以被认为是一种在热力学第二定律最终胜利之前，使生物体有足够多的热量进行繁衍的策略。非洲象代表的是一种在非洲丛林环境中获取热力学利润并繁殖的策略，而冠水母目水母代表的是一种在深海环境中获取热力学利润和繁殖的策略。冠水母目是一种特别成功的策略，因为它在大约两亿年的时间里一直是盈利的。

当然，为了获取能量和创造秩序所需材料的竞争非常激烈。植物会争夺土地、水和阳光，许多物种都有通过食用其他物种获取能源和材料的策略。生物进化是一个长达30亿年的探索，它在竞争和不断变化的世界中寻求在热力上盈利的策略。

在《熵定律与经济过程》中，乔治斯库-罗根提出，正如在生物系统中那样，“经济过程实质上是由高熵向低熵的转变构成的”。[8]乔治斯库-罗根指责传统经济学忽视了熵在经济学中的作用，并声称新古典主义经济学理论违反了物理学定律，因为它的模型中没有承认热力学的限制。[9]他认为这类似于“忽略现实世界和伊甸园之间的差异”，并将新古典主义生产函数称为“魔术”。[10]

乔治斯库-罗根的批评从来没有得到经济体制的回应，他的书陷入了半无名的状态。[11]罗根警告说，经济中的熵减少的副产品是污染。他的想法是在环境运动的基础上提出的，这也导致今天大多数经济学家将他归类为环境经济学家，忽略了他对传统理论的整个大厦提出的重大挑战。[12]

也许乔治斯库·罗根被忽视的一个原因是，熵这个概念在经济学上有一段惨淡的历史。几十年来，许多研究人员一直在经济学中寻找熵和能量的隐喻等价物，例如货币和熵之间的等价关系，或者将预算约束比作能量保护。[13]正如保罗·萨缪尔森在1970年的诺贝尔奖获奖演讲中说：“我审阅了很多篇枯燥的论文，这些论文的作者都在寻找与熵或另一种能量形式对应的东西。”[14]在1972年的一篇论文中，他进一步说道：“社会科学中的怪人和不成熟投机者的标志是，他在社会系统中寻找与物理学中‘熵’相对应的东西。”[15]

萨缪尔森是正确的。正如我们所讨论的，这种隐喻的比较就像废话。然而，这正是瓦尔拉斯、杰文斯和其他传统经济学理论的创始人所做的事情。但是乔治斯库-罗根不是怪人，他也没有在经济学中寻找与熵类似的东西。相反，罗根认为经济系统存在于真实的物理世界中。

经济系统必须遵守熵定律，像宇宙中的其他一切事物那样。[16]正如英国天体物理学家阿瑟·艾丁顿（Arthur Eddington）爵士曾发表过的著名言论：“如果你的理论违背了热力学第二定律，我无法给予你任何希望，它只能在最深重的耻辱中崩溃。”[17]

创造价值的三个条件

乔治斯库-罗根有三个重要的观察，这些观察从根本上把经济视作一个不断发展的复杂系统，可以启发我们将其与财富的起源联系起来。

第一，创造经济价值的过程本身是不可逆的。在经济系统中，时间是单向的，或者正如罗根所说，我们不能用同一块煤来驱动机车两次。[18]

第二，“现在偶然的观察都能证明我们整个的经济生活都依赖于低熵，才智、布料、木材、瓷器、铜……这些都是高度有序的结构”。[19]正如前面提到的，经济过程都是使用能量将秩序较低的原材料和信息转换成高度有序的产品和服务。

第三，尽管产品和服务本质上是一种创造秩序的活动，但并非所有的秩序都有经济价值。正如乔治斯库-罗根含糊地说道：“没有人能利用毒蘑菇的低熵，也并不是所有人都会追求海藻或甲虫的熵。”[20]

如果能同时满足以下三个条件，由物质、能量和信息构成的模式就具有经济价值。

1．不可逆性。所有创造价值的经济转换和交易在热力学上都是不可逆的。

2．熵。所有创造价值的经济转换和交易都会使经济系统内局部的熵减少，而使全球范围内的熵增加。

3．适合度。所有创造价值的经济转换和交易都会产生符合人类目的的人工制品或行为。

不可逆性：打破鸡蛋做蛋饼

早些时候，我们把企业定义为为了盈利而将物质、能量和信息从一种状态转换为另一种状态的人或团体组织。同样，我们也从转换的角度定义了物理技术和社会技术。把事物从一种状态转换为另一种状态的概念是热力学概念。[21] 19世纪的法国工程师萨迪·卡诺把所有的转换分为两类：可逆的和不可逆的。绕太阳运行的地球是一个可逆转换的例子（我们可以称之为转换，因为它的位置会随着时间的推移从一个点转换到另一个点）。地球绕轨道运行的方向没有任何特殊的原因，牛顿方程中也没有任何东西说明它不能朝着另一个方向运行。如果看到一部地球绕着一个方向旋转，然后又绕着另一个方向旋转的电影，我们将无法分辨哪个方向是前进的，哪个方向是倒退的，因为地球位置的变换在时间上是

完全对称的。[22]相反，如果看到一个牛奶瓶从桌子上掉下来砸碎在地上的电影，然后再看倒放，我们就很容易区分顺序和倒序，因为打碎牛奶瓶是一个不可逆的过程。

在生活以及经济现象的宏观层面上，我们可以参考的一件事是时间只朝一个方向运行。[23]正是热力学第二定律给了时间以方向。我们的大脑凭直觉就可以知道熵增加了，减少熵的唯一方法就是能量输入和工作。因此，如果看到有关打碎牛奶瓶和一滴墨水在水中扩散的电影，我们就会看到这些系统的秩序在下降，知道时间在流逝。如果看到破碎的牛奶瓶自发地跳起来恢复完整，或者扩散的墨水重新组成一滴有序的墨水，我们就会知道时间在倒流——这样的秩序创造不会自发进行。但是如果我们看到一个人小心地把奶瓶黏在一起并重新灌满它，我们会相信时间又向前推进了，因为有人把能量投入系统来创造秩序。正如物理学家理查德·费曼曾经指出的，如果一部关于转换的电影按倒序放映，观众大笑，那么这个过程就是不可逆转的。[24]

不可逆性与熵和秩序创造是紧密相连的，这种联系是通过概率定律产生的。思考秩序的一个有效方法是计算随机运动的分子出现某种特定状态的概率。如果你把牛奶倒进一杯咖啡，静置一会儿，不去搅拌它，牛奶最终会消散在咖啡中，因为组成牛奶和咖啡的分子随意运动，相互碰撞，使饮料达到了一个均衡状态。几分钟后，再看咖啡杯，你就会惊讶地看到黑色的咖啡聚在杯子的一边，而白色的牛奶在另一边。理论上，咖啡分子和牛奶分子的随机碰撞可能纯粹是偶然的结果，但这是一种“在宇宙存在期内”不会出现的现象。现在想象一下，我们在这杯牛奶咖啡的混合物中间设置了一个屏障，这个屏障有一个分子大小的活板门，由一个小型纳米机器人操控。通过编程，纳米机器人可以对咖啡分子和牛奶分子进行分类。每当一个随机移动的牛奶分子撞到门的一侧，机器就会打开门，让牛奶分子通过。每当机器人感觉到咖啡分子撞击门的另一侧时，它就会朝另一个方向打开门，让咖啡分子通过。随着时间的推移，机器人就将咖啡和牛奶引导向杯子的两侧。实际上，纳米机器人是在通过创造一个有史以来出现概率最低的状态来提升咖啡杯中的秩序（减少熵）。这种秩序的形成不可避免地要付出代价，我们必须给纳米机器人提供能量来完成工作，反过来，纳米机器人也会释放热量。在咖啡杯内部，局部的熵会减少，但在它周围更广阔的宇宙中，熵仍会增加。在物理学中，这样一个虚构的分子分类装置被称为麦克斯韦妖。詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）在1876年提出了这一思想实验，该装置就是以他的名字命名的。[25]经历了大约140年的讨论、提

炼和实验，科学家们得出结论，没有某种能量来源的话，麦克斯韦妖是不能使熵逆转的——没有免费的秩序。

乔治斯库-罗根说，如果宇宙不能摆脱热力学第二定律，那么经济学也不能。[26]创造经济价值的转换或其他过程在热力学上来说是不可逆的。这并不意味着逆转价值创造过程是不可能的，相反，它意味着制造事物需要能量，使事物恢复原状也需要能量。在经济体系中，时间是有方向的。例如，造纸的过程就包括砍倒森林里的树木，装上卡车，木材被制成纸浆，与化学物质混合制成浆料，然后把浆料压成卷，干燥，切成薄片。而用一堆纸还原出一棵树是非常困难的。使用费曼的测试，如果观看一部纸浆变成木材，木材从卡车上跳下来的电影，我们马上就会知道它在倒放。同样，汽车自我拆解，飞机向后倒，银行出纳员在电脑上打的字消失等电影都会让观众发笑。

想象一种完全可以逆转的经济转换过程是可能的。1982年，IBM公司的物理学家查尔斯·贝内特（Charles Bennett）展示了可逆计算机在理论上的可能性，即一台既可以向前又可以向后执行计算而不产生任何热力损失的机器。如果这样的计算机存在的话，就可以完成有益经济的事，这似乎违背了乔治斯库-罗根提出的第一个条件。但是接下来，贝内特在IBM的同事表示，这样一台可逆的计算机需要无限的记忆存储。一旦计算机开始清除记忆中的信息，为新信息留出空间，它就有一个不可逆的过程。只要经济是有限的（例如，经济是宇宙的一个子系统，而不是宇宙本身），那么乔治斯库-罗根的理论就是安全的，也没有违背他提出的经济不可逆的条件。[27]

乔治斯库-罗根的第一个条件表明，所有具有经济价值的产品和服务都是由热动力学不可逆变换产生的。简单地说，你不可能不打碎鸡蛋就做出蛋饼。

你可能已经注意到，在不可逆条件的措辞中，我将交易一词加到了转换的后面。这是因为不仅生产商品和提供服务能够创造价值，而且交易也能创造价值。正如本书前面所讨论的，传统经济学早就认识到了这一点，这是由人们的不同偏好而产生的，交易能按照人们的偏好配置世界上的商品和服务，从而创造价值。正如创造价值的生产是不可逆转的，创造价值的交易也是不可逆转的。[28]从直观的角度来说，如果两个人因为互利而同意进行交易，那么他们就不会取消，在之后马上是这一过程的逆转。这样做会使他们每个人都失去交易的价值（当然，除非新信息——如交易货物有缺陷的信息被揭露，导致其中一个人改变主

意）。

从物理和热力学的角度来看，描述交易的特征有些复杂。例如，如果看到人们用汉克·阿伦（Hank　　Aaron）的棒球卡换一张“贝比”鲁斯（“Babe”Ruth）的棒球卡，我们就无法判断交易是向前进行还是向后退——是鲁斯换阿伦还是阿伦换鲁斯？这是因为我们需要知道一些隐藏信息，比如每个交易者的偏好以及他们现有的棒球卡片。假设有一位市场研究员在交易前调查两位交易者，询问他们最喜欢的棒球运动员是谁，他们分别有什么卡片。然后，我们将看到在交易之后他们的偏好与现有的卡片更加匹配了，这样我们就可以知道交易进行的方向了。[29]我们还知道，这两个交易者不会立即撤销交易。因此，在热力学意义上看，这个过程是不可逆的。

不可逆性是创造价值的必要不充分条件。我们不难想象破坏价值的不可逆过程，飓风、爆炸以及无能管理团队都会破坏价值，需要能量去转变局面。乔治斯库-罗根的第一个条件给经济系统中的时间赋予了箭头，但我们需要第二个条件给熵一个箭头，并进一步缩小创造价值的转换和交易。

降低的熵：粉色的汽车和炸弹能创造价值吗

乔治斯库-罗根的第二个条件“所有创造价值的经济转换和交易都会使经济系统内局部的熵减少，全球范围内的熵增加”使我们能够对比区分将石头扔向窗户以及修窗户的过程。任何创造价值的过程都是既不可逆又会使熵降低的。

多数经济转换过程都明显地使熵降低了，正如乔治斯库-罗根有关布料、木材、瓷器和铜的例子所展示的那样。但是，经济转换就不会涉及熵的增加吗？例如，拆除建筑物不会使价值增长吗？为军队制造炸弹呢？

我认为拆除建筑物是一个更大的经济转换的中间步骤，而不是一个完整的转换，就像把树变成木浆是造纸的中间步骤那样。人们不只是为了拆掉一座建筑物而拆掉一座建筑物。[30]拆掉建筑物是为了建造其他建筑物，或者收回土地另作他用，例如修建一条公路，甚至把土地归还自然。[31]同样的论点也适用于垃圾处理、环境清理和回收利用。所有这些都可以看作创造产品或服务的更大转换过程的一部分，这也是像乔治斯库-罗根这样的环保主义者的观点：要综合地看待生产和清理。破坏作为

价值创造的中间步骤在生物系统中的作用也是如此。身体在细胞和系统中创造新的秩序之前，你的消化系统必须分解食物中井然有序的化学物质和能量。

然而，爆炸物是另一回事。把一堆化学物质转换成炸药的行为是一种降低熵的活动。爆炸物制造者利用能量和知识来获取无序分子，并将它们排列成整齐、高势能的能量包，这些能量包可用于采矿或拆除等工作。从热力学的角度来看，分子的随机运动极不可能自发地把自己排列成一根炸药。当军队使用炸药炸毁敌方目标时，它可能是在做一些保卫国家安全的重要事情，但没有人会认为使用炸弹是一种创造经济价值的转换。制造炸弹能使熵降低，并且创造价值（有证据表明有人愿意为此付出代价），但是使用炸弹会通过炸毁敌人的资产而使熵增加，这一行为会破坏而不是创造经济价值。战争不可避免地会增加熵，因为它将来之不易的经济秩序变成了废墟，将人体的生物秩序变成了伤害和死亡。

到目前为止，我使用“秩序”一词都非常随意，就像法官波特·斯图尔特对“色情”的定义那样，“我们看到它就会知道”。然而，秩序是一个模糊的概念。例如，我可以在半夜溜出去，把密歇根州安阿伯的每辆车都漆成亮粉色。[32]很难说这种做法会创造价值，毕竟安阿伯的居民可能会对我很生气，也不可能为我的努力付钱。不过，我可以说我已经使熵降低了。毕竟，我花费精力创造了一个概率更低的世界格局，一个所有汽车都是粉色的城市比一个汽车颜色混杂的城市更不容易出现。但你也可以说，这个活动实际上是使熵增加的。通过把每辆车都刷成粉色，我破坏了信息——在粉刷之后，人们很难辨认出自己的车。尽管我费力做了这件事情，但它并不比用石头砸窗户或者其他破坏行为更好。那么，把汽车刷成粉色是创造了秩序还是破坏了信息呢？

答案取决于你的观点，这是理解秩序这一概念的关键点。秩序和无序的构成都需要与某事关联起来去衡量。对一个在直升机里努力计算安阿伯汽车数量的交通规划师来说，把所有的车涂成粉色可能是在创造秩序，因为这能帮助她更准确地定位汽车，她甚至可能愿意付钱让我去做这件事。但从车主的角度来说，这是在破坏秩序，因为关于车的颜色的信息丢失了。至于车主更喜欢哪种颜色，这就是另一回事了。

这种秩序的相对性在热力学中是一个众所周知的问题。波兰物理学家沃伊切赫·楚雷克（Wojciech Zurek）举了一个他称作“洗牌把戏”的例子。[33]一个魔术师向你展示了一副牌，牌面完全按照数值和花色排列（两颗红心之后是三颗红心，以此类推，另外三种花色也是如此）。魔

术师完全随机地把牌彻底洗了一遍，然后他向你展示这个随机的牌面。接下来，他给你另一副牌，同样按照数值和花色完美排列，然后用20美元打赌，说你不能完成他刚刚做的事。你想着洗牌并不难，于是接受赌注，好好洗了洗这副排列有序的牌，直到它看起来非常随机。你微笑着把它还给魔术师，等待着你的20美元。但是他说："你输了，我洗过的牌的顺序是黑桃皇后、梅花10、方块4……你的牌和我的根本不一样，你欠我20美元！"这个寓言的重点是，牌面的任何一种顺序都可能是也可能不是其他的顺序——把牌洗成完全按照数值和花色完美排列的顺序和其他特定顺序的概率是相同的。我们只能从具有相同可能性的顺序中挑出特定的一种，然后称之为"有序的"。

低熵可能确实是某事物具有经济价值的必要条件，但要定义哪种秩序是有价值的，哪种是没有价值的似乎非常主观——秩序是个人的主观感受。因此，乔治斯库-罗根的前两个条件对于创造价值都是必要但不充分的。

适合度（第一部分）：偏好的进化观

不可逆性和秩序从根本上讲是与创造经济价值相连的，但我们需要理解为什么与其他秩序相比，人类更喜欢某种秩序。为什么我们更喜欢具有独特颜色的车，更喜欢脆苹果，而不是脆甲虫呢？一种探索方法是走传统经济学的道路，简单地假设人类有偏好，这些偏好以合乎逻辑的方式排列，人们以最大限度满足这些偏好的方式行动。在这种情况下，我们不需要知道这些偏好是什么——它们在人们进行交易和消费时会显现出来，我们不需要知道它们是如何形成的，或者它们会如何随着时间而改变。按照这种观点，财富只不过是"人们偏爱的任何秩序模式"。但这也有些令人不满，因为很多行为都被转移到了一个叫做偏好的神秘的黑匣子中。让我们看看能否打开这个盒子，把它与我们正在构建的更大的进化系统相连。[34]我们为什么想要自己想要的东西呢？你为什么想要一顿美味的晚餐，一套新衣服，一次热带岛屿之旅，一辆红色跑车或者一条钻石项链呢？你为什么更喜欢钻石项链而不是美味的晚餐呢？

偏好是一种心理现象。弗洛伊德假设，人类的物质需求是由动物本能驱动的，并不断地由我们的超我进行检视，因此，我们的经济偏好是在"我真的想要那辆昂贵的车！"和"但我需要为孩子的教育存钱！"这两者间斗争的结果。弗雷德里克·斯金纳（B. F. Skinner）认为，偏好基本

上是后天习得的。一个斯金纳理论的支持者可能会说我们想要一辆豪车，是因为我们从社会中了解到豪车是值得拥有的，汽车制造商会通过展示诱人的营销信息让我们买它。介于弗洛伊德的动物本能驱动理论与斯金纳的习得行为之间的理论是由心理学专家亚伯拉罕·马斯洛（Abraham Maslow）在20世纪60年代提出的。马斯洛声称人类有一个“需要层次”，从食物、水、性、住所、睡眠等基本的身体需求开始，然后进入更高层次的需求，如自尊和社会尊重。一个已经满足了基本食物和住房需求的中产阶级可能会在豪车或者衣服上大肆挥霍，从而满足他对社会尊重的需求。马斯洛需求层次理论的最高层是自我实现。当你成为亿万富翁，拥有了一切，你就可以把物质主义放在一边，去做公益，去西藏寻找自己，或者坐在浴缸里思考生活的意义。

尽管马斯洛用一个有用的（但非常有争议的）框架组织划分了我们的需求，但他的理论并没有回答更深层次的问题，比如这些需求来自哪里，为什么作为消费者，我们更喜欢某些东西而不是其他东西。这些更深层次的问题的答案可能来自进化心理学最近的研究。[35]进化心理学认为，我们的基因构建大脑时有一个目的：让基因遗传给下一代。人们常常把进化心理学与“所有行为都是由基因决定的”这一观点混淆。任何进化心理学家都不会说某人对汉堡王的偏好高于麦当劳是因为他的基因。相反，进化心理学家所说的是，我们的许多行为都是这样的，因为它们帮助了我们的祖先在距今50万年前到10万年前的非洲大草原上生存繁衍。尽管所有人类都有共同的行为，例如对食物、性和社会地位的渴望，但人类也有很强的适应能力，正如罗伯特·赖特所说，这些行为已经适应了当地的环境。[36]对于进化心理学家来说，“先天还是后天”并不是一个有趣的问题，显然这两个因素都能决定行为。相反，有趣的问题是，为什么“先天”这部分进化了，环境如何影响和调整“后天”这部分，以及人类行为的这两面是如何相互作用的？

从人类学的角度观察狩猎采集部落，我们就可以看到偏好行为如何与帮助人们在自然环境中生存、竞争、繁衍及养育后代等直接联系在一起。我们很难从抽象的现代生活中看到这种联系，但一位进化心理学家会说，潜伏在现代经济偏好之下的，是人类在非洲大草原上的狩猎采集生活方式进化的影子。据美国劳工统计局统计，美国人90%的消费支出可以分为七大类。[37]如果从进化心理学的角度来看这些类别，就会发现一种模式：

◎ 住房（占总支出的32%）。我们倾向于看重那些能在恶劣

天气里保护我们的，为家庭成员提供良好环境的住房，并希望通过房子的大小、装修方式和所在位置来显示自己的社会地位。能保护我们、有利于抚养孩子的住所显然是具有进化价值的。用住所衡量社会地位和财富可以一直追溯到人类定居的考古学证据。[38]麻省理工学院认知科学家史蒂芬·平克（Steven Pinker）(15)甚至说，现代人类对具有吸引力的物理环境的一些偏好，建立在人类早期历史中对成功进化的环境的界定。[39]

- 交通工具（占总支出的20%）。交通工具既能帮助我们谋生，又能增强我们与亲友的联系。男性有时也会通过交通工具来彰显自己的社会地位并（试图）吸引女性。

- 食物（占总支出的14%）。我们会珍视某些食物的独特口味，特别是鲜类、香甜和高脂肪含量的食物，因为这些食物在人类祖先的生存环境中非常稀缺并具有较高的营养价值。[40]正如罗伯特·赖特所说，许多企业，从麦当劳到高档法国餐馆，都是通过满足这些味道来经营的，在每个街角都能提供饱和脂肪以及多数人都坐着办公之前这些味道就已经进化出来了。[41]

- 人寿保险和养老金（占总支出的9%）。如果你不幸早亡，人寿保险能使你的后代及伴侣生存下去的可能性达到最大。养老金能够确保你在完成生育并且孩子已经长大后，你不用担心进化价值的降低仍能活下去。

- 医疗保健（占总支出的5%）。良好的医疗保健的进化益处是显而易见的，正如进化论所预测的，我们往往最关心自己和近亲的身体健康。

- 服装（占总支出的5%）。除了通过衣物在恶劣天气里保护我们外，人们还倾向于通过衣物来彰显身份，表明自己是某个特定群体的成员以及吸引伴侣。卡地亚手表可能会彰显你是一个地位显赫的人，你在鞋子上的选择可以表明你是嘻哈歌手还是税务会计团队的成员，而性感的无肩带礼服或时髦的西装肯定会帮助你吸引潜在的伴侣。更实用的品牌也能满足同样的需求，但传达的信息不同。例如，一个穿着Lands'End品牌的男性可能是在发出这样的信号："我是一个穿着斜纹棉布蓝色衬衫的理性中产阶级，我

是一个可靠的伴侣，不喜欢炫耀。”

◎ 娱乐、媒体和通信（占总支出的5%）。乐趣的概念也与进化的益处有关，比如和朋友待在一起，通过晚餐建立联系或者在酒吧里寻找伴侣。运动，以及一些为了争夺地位和资金的战争都能够追溯到原始人生活的时代。同样，故事、新闻、八卦和一般信息在人类祖先生存的环境中也带来了巨大的生存效益。一项关于手机使用情况的研究表明，现代人用手机建立和维护社交网络的方式与其他灵长类动物用身体装束建立和维护社交网络的方式大致相同。[42]另一项研究发现，相比于独处或在其他男性面前，男性手机用户更倾向于在女性面前掏出手机夸夸其谈，从而用手机来彰显财富、地位和社会联系，所有这些都是在人类祖先中提升繁殖成功机会的因素。[43]

适合度（第二部分）：按下我们自己的快乐按钮

巨无霸、保时捷、Jimmy Choo鞋和手机在早期的非洲大草原上根本不存在。而许多在那个环境中进化出来的能够增加人类生存繁殖可能性的驱动力，仍体现在我们今天的需要、欲望以及情感上，并且能解释为什么人类在现代消费社会中想得到我们想要的东西。

尽管在食物、住房、衣服、医疗保健甚至手机上花费巨大可能具有进化逻辑，但我们在某些事情上的花费真的很难看到与进化有什么关系。例如，我们为什么会想花钱买一幅画或者听一首曲子？愤世嫉俗的人可能会说这也与象征地位或者吸引配偶的企图有关。正如平克所说，“没有什么比在小玩意儿上花钱更能展示你有闲钱了……它们既不能填饱肚子又不能遮风挡雨”。[44]纵观历史，喜欢及支持艺术的多是富人。然而，许多人会说，审美体验的乐趣本身就具有价值，并不是只有富人才会享受这些。那么在人类祖先生活的环境中，对艺术的渴望是如何与进化优势联系在一起的呢？答案是，两者并不相关。根据进化心理学家的说法，艺术就是进化论中的功能变异。我们可能由于其他原因进化出了关于什么能构成美丽的风景、有吸引力的脸庞、愉快的声音的想法，例如，世界各地的人们都会被水、高地、树荫等景观所吸引，这些景观可能与人类祖先生存的环境所打造出来审美有很大的关系。同样，被视为美丽的脸往往是对称的，这是年轻和生殖健康的标志。

平克把艺术称为“精神奶酪蛋糕”，[45]奶酪蛋糕在远古环境中并不存在，但很多人都对它有着渴望。人们对奶酪蛋糕的渴望源于对脂肪和糖的渴求，这两者在早期都是很稀少的。此外，就像后来学会如何制作奶酪蛋糕以满足早期的生理渴望那样，我们也学会了生产艺术品来满足早期对美好图像和声音的心理渴望。这一切都不能否认人们在体验艺术时感受到的强烈且真实的情感，也不能否认创造艺术的天才的独特性。进化心理学仅仅为起初艺术为什么会存在提供了一些解释。

人类心理在早期环境中进化出一系列快乐按钮和痛苦按钮，因为这些按钮可以帮助人类生存繁殖。拥有大智慧和灵巧双手的人类已经想出了如何用各种巧妙的方法按下自己的精神愉悦按钮，从麦当劳到高档美食，从色情到艺术。传统经济学倾向于关注偏好的差异，但更有趣的问题是，在一个拥有70亿人口的世界里，为什么人类的偏好有如此多的相似之处？为什么全世界的人都喜欢像可口可乐这样的碳酸饮料？这让营养学家和反全球化的抗议者非常懊恼。为什么从俄罗斯到巴西的青少年都渴望最新的耐克运动鞋？为什么电视节目《海滩游侠》（*Bay Watch*）在美国郊区和中国农村都大受欢迎？对于进化心理学家来说，答案是显而易见的——含糖饮料、能彰显地位的物品以及性欲旺盛的人们按下了很久以前进化出来的精神愉悦按钮。虽然进化心理学在研究经济偏好和消费者行为方面的应用仍然是新兴的，并带有推测性质，但它可以让我们深入了解为什么某些形式的秩序对人类而言比其他秩序更有吸引力。

在经济发展框架中，我们可以看到，人类不断变化的偏好为商业计划的进化提供了适合度的约束，商业计划和消费者偏好是共同进化的。同理，这就是进化论者所说的生态位构建：生物体在进化中对环境产生影响，环境又反过来影响着生物体的进化。[46]例如，植物吸收二氧化碳并释放氧气，而需氧生物（如蠕虫、鱼、哺乳动物）会吸收氧气并释放二氧化碳。这两组生物体随着时间的推移共同进化，将大气中的氧气含量从1%提升到了21%，这反过来又为未来的进化提供了适合度的限制。

在经济领域，人们的需求和品位是共同进化的，而商业计划则通过进化来满足这两者。例如，灵长类动物的听觉进化是一种生存机制，能很好地在非洲大草原的环境中识别出重要的声音模式，例如猎物或食肉动物出现的声音。随着人类语言能力的提高，听觉进一步进化。大约3万年前，音乐发展为某种“听觉奶酪蛋糕”，来满足人类不断进化的听觉和心理。[47]对MP3播放器的渴望与我们的生存和繁殖并没有直接的关

联。在MP3被发明之前，我们也无法渴望拥有它。基于上述原因，人类形成了对音乐的爱好，自那以后，商业计划就被制定出来以满足需求——从制作并交易骨笛和皮鼓，到销售吉他、口琴以及后来的晶体管收音机和MP3播放器。我们对音乐的渴望由来已久，背后有着深刻的进化逻辑，随着时间的推移，满足这种渴望的解决方案也发生了重大变化。人类的偏好推动了商业计划的进化，而商业计划的进化也影响着人类偏好的进化。

通用效用函数

总体来看，人们的需求、品位和商业计划表明经济活动本质上是关于秩序创造的。面对着世界的混乱和随机性，人类醒着的多数时间都在以不同的方式组织环境，使其成为一个宜人、舒适的地方。人们通过将能源、物质和信息转化为想要的商品和服务来组织世界，通过合作、专业化和贸易发现了进化的把戏，通过这些把戏我们创造了自身无法创造的秩序。但人类匆匆忙忙、鬼鬼祟祟创造秩序的目的是什么呢？我们为什么会这么做？

在这一点上，乔治斯库-罗根提出，我们所有创造秩序的活动都是为了增加人类的幸福感。他给出的解释追溯到了边沁关于效用的概念——有一种神秘的以太，也就是乔治斯库-罗根所说的心流，它不断地在人身上流动，能够即时测量人在某个时间点的瞬时幸福感。[48]从进化的角度看，乔治斯库·罗根说所有熵降低的活动的适合度函数都是人类的个人幸福。

接下来，我们将告别乔治斯库-罗根，因为我看到了一个不那么神秘且更科学的选择。生物进化理论中只有一个普遍的效用函数——基因的复制。从基因的视角来看，基因构建身体是一种自我复制的策略。这个策略的一部分是给身体提供大脑，使人类能够生活在复杂、合作的社会环境中，并制作工具。建造这些大脑是为了复制制造它们的基因，结果，大脑形成了目标、偏好和欲望，这些与人类祖先的生存、交配以及养育后代是一致的。我们利用大脑来调节环境，以满足自己的目标、偏好和欲望。

理查德·道金斯明确地指出，为历史的基因复制利益服务与为我们现在的幸福利益服务是不同的。[49]基因让我们渴望先前环境中缺乏的脂肪和糖，但在现代世界，这些物质对我们的伤害大于好处。同样，基因给

了我们情绪上潜在的非理性愤怒发作。在语言发展前的狩猎采集时代，惩罚是合作等式的另一半，沟通的作用是有限的。但在现代社会，有了法律、规范和其他强制机制，这种不合理的爆发带来的更多是伤害而不是好处。进化论并没有给我们带来幸福，也不一定能带来让我们进入更幸福的状态的目标、偏好和欲望。进化赋予人类的就是在过去恰好能满足我们的生存繁殖利益的策略。从逻辑上讲，这就是它能给我们的一切。

尽管进化的精神启示超出了这本书的范围，但它确实解释了现代生活的一个谜：为什么钱买不到幸福？这不只是老生常谈，而是一个经验事实。许多心理学家，包括诺贝尔奖获得者、普林斯顿大学的丹尼尔·卡尼曼，都仔细研究了人类在不同文化、不同时代中感到幸福的原因。[50] 他们发现了一个强有力的直接基因联系，可以解释大约50%的幸福感。鉴于科学家们知道大脑中有关于幸福的强大生化基础（这也是抗抑郁药物对许多人有效的原因），因此遗传学对个体大脑化学组成的相对幸福有影响也就不奇怪了。卡尼曼和他的同事发现，婚姻、社会关系、就业、社会地位和身体等因素都对幸福感有显著的影响。另外，从进化的角度看，拥有伴侣、社会联系、较高的社会地位和舒适的居住环境会触发大脑中与“快乐”有关的化学物质——这些都是祖先环境中基因成功复制的关键因素。

一个人拥有的财富的绝对水平确实会对其幸福感造成影响，但这种影响并不是线性的。极度贫困、需要为生存斗争的人往往会不太开心。然而，一旦人们达到了能满足基本需求的经济水平，财富和幸福感的关联就会稳定下来。过了这个门槛后，人们倾向于以相对而非绝对的基础来看待财富。财富的增加，特别是意想不到的财富的增加，会让我们感到快乐，但是我们很快就会回到幸福基线。意外的加薪可能会让你开心几个月，但之后你就会习惯，一段时间后，你又会觉得自己所得报酬过低。对彩票中奖者的研究也显示出相似的模式，他们最初会体验到巨大的快乐，期望生活中的问题能得以解决，但之后又会退回幸福基线，甚至是更低水平。同样，这种对待财富的态度是有进化意义的。可以想象，在一个竞争激烈的世界里，不断奋斗、积累财富、一刻不停、永不满足的基因是如何击败饱足、肥胖且快乐的基因的。贪婪可能不利于我们自己或他人的幸福，但一定数量的贪婪在历史上有利于人类基因的复制。

财富就是“适合的秩序”

所有的财富都是由热力学上不可逆的熵降低过程创造的。创造财富的行为可以创造秩序，但并非所有的秩序都能创造财富。个人的头脑、组织和市场在商业计划中发挥作用，寻求不同形式的经济秩序，而这些经济秩序是由市场计划、消费者安排的。消费者们选择能满足他们的需求、欲望及偏好的秩序形式，这看起来很现代，却在人类基因的通用效用函数中有着深厚的历史根源。因此，财富是一种反熵的形式。它是秩序的一种形式，但并不是任何秩序，只是适合的秩序。经济秩序的模式以产品和服务的形式相互竞争，以满足消费者的需要、期望甚至渴望。在竞争中成功满足人类偏好的经济秩序模式是适合的，并且有助于创建适合经济秩序的商业计划模块会随着时间的推移而被放大。正如物种和环境是共同进化的，商业计划的竞争生态系统和消费者的偏好也是共同进化的，这会使适合度成为一个有条件的概念，今天适合的秩序明天可能适合也可能不适合。

我们可以将财富作为“适应的秩序”的理念与传统的经济价值概念联系起来。在古典时期，经济学家认为经济中的价值来源在于供给侧，这种价值来自生产要素。例如，坎蒂隆认为价值是生产产品所用的稀缺土地数量的函数，马克思认为劳动是价值的最终来源，李嘉图补充说资本也是必要的。有了杰文斯和边际主义者，价值理论转向了需求侧，因为他们认为价值是由商品的相对效用决定的。在新古典主义的综合推理中，两者是融合在一起的：稀缺的生产要素通过市场机制满足了个体的偏好，价值只是两个人愿意交易的任何东西。如果小部件制造商愿意以一美元的价格出售小部件，而消费者愿意以一美元的价格购买小部件，那么小部件的价值就是一美元。

价值的进化观包括供给侧和需求侧。在供给侧，低熵的东西是有价值的。根据定义，低熵的事物是稀缺的，需要能量、材料和信息来创造它们。在需求侧，人们的偏好决定了相互竞争吸引我们注意力的产品和服务的相对吸引力。在传统经济学中，双方通过市场机制相遇，寻找个人偏好与稀缺秩序相匹配的商业计划。金钱是一种社会技术，它能使我们在这些互动中保持得分。在复杂经济学中，一个小部件的价值仍然是一美元，但我们可以更深入地理解其中的缘由。

经济财富和生物财富在热力学上属于同一种现象，而不仅仅是比喻。两者都是局部低熵系统，是在适合度函数的约束下随时间进化的秩

序模式。经济的适合度函数，即我们的品位和偏好与生物学的适合度函数有着根本的联系——基因的复制。经济本质上是一种基因复制策略，这是另一个进化的把戏，还有猎豹的伪装、蝙蝠的雷达和果蝇的眼睛。经济是一个非常复杂的把戏，建立在强大的智慧、灵活的制造工具的手、合作的本能、语言和文化之上。[51]

如果财富确实是适合的秩序，那么我们可以用另一个更熟悉的词去描述它。在物理学中，秩序和信息是一样的，因此我们也可以把财富看作适合的信息，也可以称其为知识。一方面，信息本身可能毫无价值；另一方面，知识又是有用的，比如那些我们可以利用的、满足某些目的的信息。因此，我们就有了一个完整的循环，传统增长理论的创始人罗伯特·索洛是对的——财富的来源是知识。然而，我所概述的基于复杂性的观点并没有将知识视为一种假设、一种外源输入、一个超出经济学界限的神秘过程，而是将知识的创造性置于经济的内在核心。[52]

进化是一种创造知识的机器，是一种学习算法。[53]生物世界中有许多巧妙设计中蕴含着丰富的知识。蚱蜢是工程学的奇迹，是物理、化学和生物力学知识的仓库——这些知识超出了目前人类复制能力的范围。蚱蜢也是关于其进化环境、食物、需要防御的掠食者以及在吸引配偶和确保后代存活方面起到良好作用的策略的集合。一只蚱蜢中蕴含着数兆字节的知识，而整个生物圈之内更是蕴含着惊人的大量知识。所有的秩序、复杂性以及知识都是由最简单的方法创建并聚集的：变异、选择和放大。

现在，环顾你所在的房间，想想你周围的物体中蕴含的全部知识。椅子制作中蕴含的木工手艺；服装中蕴含的棉花种植、织物制作、时尚设计知识；电灯中蕴含的电力及材料知识；书本中蕴含的各学科知识。经济圈中蕴含的知识总量和生物圈中蕴含的知识总量一样惊人，经济圈也是通过变异、选择和放大创造出来的。

我们已经找到了长久以来探寻的答案：财富是知识，它的起源是进化。

我们通过测试了吗

在本书的开头，我描述了一个任何要解释财富创造的理论都必须通过的测试。该理论必须能从一个由一群人和一组自然资源组成的“自然状态”开始，创造一段熵不断减少，复杂性、组织性、多样性和财富不断增

长的历史。这个理论必须把我们从制造石器工具的原始人变成畅饮夏敦埃酒的纽约人，并在历史记录中不连续的、爆发性的、间断性均衡的模式中也这么做。此外，该理论必须用最低限度的特别假设做到这一切，而不必为关键动力触及边界之外。这一理论必须与科学界其他公认的理论保持一致，而不是相互矛盾。

这是一个很高的秩序，现在这样的理论还不完全存在，但我相信它是可实现的。在之前的5章中，我把许多人的作品综合成了一个像这样的理论。大约250万年前，我们可以从本质上“启动”人类历史上的模式，以发明的第一把石斧作为物理技术空间的起点，以近亲的小团体作为社会技术空间的起点。仅有的外源因素是能量和物质的物理输入，以及热量和废物的输出。仅有的外源动力是与生物学相关的，例如人类可以演绎-探寻的大脑的生长、语言的发展和基本偏好的进化。这些一旦建立起来，通过物理技术、社会技术和商业计划空间进行的搜索将被关闭起来运行，所有空间相互依赖、共同合作，伴随着连续的发现呈指数级扩展。在这三个设计空间之外，我们会看到不断发展的模式，充满快速变化和缓慢变化的间断性阶段，同时快速地朝着更低的熵、更高的复杂性和多样性以及更多的财富发展。值得注意的是，这种走向更好的秩序和更多的财富的趋势并不是进化的必然结果；适合的条件是必备的。此外，我们也无法保证重新“播放”进化的录像带（用史蒂芬·杰伊·古尔德的话说），还能实现与今天这样的水平相当的经济。进化是由数百万微小事故和偶然转折积累起来的，即使是历史上的一个微小变化也可能导致截然不同的结果。[54]但这种理论将有助于我们了解经济复杂性出现和发展所需的条件，以及促成今天经济发展的广泛参数。

乔治斯库-罗根的深刻见解都体现在了他的三个条件中，这些条件为类似的理论提供了一个潜在的重要基点，事实上，人们可能会说前两个条件，即不可逆性和降低的熵构成了经济学的经验定律。总的来说，当某物成为自然界中的普遍规律，并且没有已知的例外时，它就能被称作定律。[55]定律本身不一定能解释规律存在的原因，它只是指出规律确实存在。理论能够解释规律为什么存在以及它是如何运作的。当我们既有经验定律，又有能深刻解释这些法则的理论时，科学就达到了顶峰。例如，热力学定律本身就是经验规律，但科学对于为什么这些规律会存在有着深厚的量子级的理解。相反，正如我们在第3章中看到的，仔细审视就会发现经济学中的定律通常是近似的。乔治斯库-罗根的前两个条件可以被视为定律，因为它们的经验是可测试的，而且我会说这些定律是正确且没有例外的。乔治斯库-罗根的第三个条件就有些复杂了，因为我

们已经注意到，适合度只能在事后观察到，并且我们观察经济适合度的唯一方法就是看人们是否愿意为某事付款或进行交易。因此，乔治斯库-罗根的第三个条件更像是一个逻辑结构，而不是经验定律。

在建立一个严格的、数学的、经验上可检验的经济进化模型来解决所有这些问题之前，我们还有许多研究要做。但我在第二部分和第三部分中描述的复杂经济学中的广泛运作体表明，拼图中的许多模块已经就位，这样的理论不仅是可能的，而且很有可能在未来几年出现。毫无疑问，这些章节中讨论的一些猜想会被证明是错误的，或者需要修改，但其中有一条清晰的线索，将我们从经济的平衡观引向进化的、基于复杂性的经济观。这种范式转移的含义是多方面的、意义深远的，它们将会成为第四部分的主题。

第四部分

财富的积累及增长：如何利用复杂经济学理论决胜于商业社会

THE ORIGIN OF WEALTH

THE ORIGIN OF WEALTH

要命令大自然，必须先顺从它。

——弗朗西斯·培根

15

战略：把机构和社会设计成更好的进化体

红桃皇后竞赛

传统经济学告诉我们，只要人的行为足够理性，并且拥有足够的信息，那么经济就会成为如钟表般可以被预测的世界。连传统经济学的不确定性都已算是中规中矩了。[1]科学界对钟表宇宙的幻想终结于20世纪，而经济学对钟表世界的幻想正在21世纪瓦解。经济过于复杂、过于非线性、过于动态，而且对于偶然的波折过于敏感，所以只有超短期的经济是可以进行预测的。[2]即使我们尽可能地保持理性，并且拥有所有需要的信息，经济的计算复杂度依然使我们来不及预测未来。[3]

这是一则令人清醒的讯息。我们常常希望人们作出需要长期预测的经济决定。经济学家们试图预测GDP或者通货膨胀的变化；首席执行官们需要确定某种商业策略是否比其他策略在未来能产生更好的结果；投资者们希望确定某种投资方式是否比其他方式能在未来产生更高的收益；而政界领袖则需要决定某种政策是否比其他政策能在未来更好地服务社会需求。

如果复杂经济学的这个讯息是正确的，我们该何去何从？人们现在都在勉强应付。首席执行官、投资者和政策制定者都在尽可能地保持着自己的理性，他们所做的预测和决定时而正确、时而错误，还要进行试验、调整方向以及持续性地修正。那么复杂经济学是在告诉我们，我们只能做到这个地步了吗？

复杂经济学打破了我们对于控制经济命运的幻想，但同时也递给了我们一个工具——一个我们一直拥有却从未珍惜的工具。我们也许无法预测或者掌控经济的发展，却可以把机构和社会设计成更好或更糟的进化体。

我们已经看到了这个工具的力量。在18世纪和19世纪，人类在科学

和市场层面发展了社会技术，导致经济飞速发展，财富以前所未有的方式聚集。复杂经济学告诉我们，发展可能会超出我们的认知，但与其试图战胜它，不如去理解它，利用它的力量服务人类的需求。在第四部分，我们将从商业战略的角度阐述应该如何实现这些目的。

你作出承诺了吗

世上有多少商学院教授和管理学大师，就有多少对商业战略的定义。[4]以下是最初的一种定义，也依然是最好的定义，由商业史学家艾尔弗雷德·钱德勒在他1962年的著作《战略与结构》（*Strategy and Structure*）中提出：商业战略可以被定义为对企业基础性长期目标所做的预设，以及为实现这些目标所制定的行动方针和资源分配方案。[5]

钱德勒的定义是一个很好的开头，因为它捕捉到了管理者们对战略进行定义时的两个关键要点。第一，战略本身是具有前瞻性的。为了制定一个战略，管理者必须决定好未来要走到哪一步。第二，管理者需要构想出一个计划，以便达到理想的未来状态，并且致力于落实计划所需要的行动方针。

这两个因素存在于大部分企业的战略核心。[6]制定计划的过程一般是从分析情况开始的，然后回顾企业和行业的发展脉络。接下来管理团队会评估影响客户、竞争对手、技术等因素的重要趋势，最后给出方案并评估这些方案中的吸引力定位。吸引力定位可以使企业拥有持续的竞争优势，比同行赚得更多。

传统经济学预测，在竞争均衡的情况下，没有哪家企业赚取的利润会高于资金供应者所需的最低收益。哈佛大学商学院教授迈克尔·波特（Michael Porter）在他20世纪80年代的开创性著作中提出，企业若想赚取更高的利润，唯一的方式就是创造竞争优势资源，以便（合法地）减少竞争并增强企业的市场地位。[7]竞争优势资源既能帮助企业降低成本，又能突出产品优势，强化专利技术、品牌影响力以及和主要合作方的特殊关系之类的因素。[8]例如，沃尔玛公司管理物流与库存的独家方法、品牌知名度以及对供应商的控制力，使它赚的利润比其他竞争对手更高。

一旦我们为了所预测的未来状态找到保持竞争优势的位置，就需要制定一个能够抵达那里的计划。另一位哈佛大学商学院的教授潘卡基·格玛沃特（Pankaj Ghemawat）发现，所有的战略计划都包含被他称

为“承诺”的东西。[9]的确，战略选择一旦作出，再撤销就很难或者说成本很高了。这类难以撤销的承诺可能包括进入一个新市场、进行一次收购以及对特定品牌的定位投资。承诺使得战略决策有别于战术决定。例如，商品降价如果很容易修改（例如你可以很容易在降价之后再把价格提高），就是一个战术决定。拿沃尔玛公司来说，长期、大范围地公开承诺“每天低价”就是一个战略决策，这样的承诺很难再反悔。1519年，西班牙征服者科尔蒂斯（Cortés）登陆墨西哥东南海岸时，命令手下焚毁船只，迫使他们征服内陆——这就是承诺。

建立持续竞争优势位置需要承诺，因为不需要承诺所能达到的位置很容易被模仿。例如，一个零售商很容易就能跟随另一个零售商暂时降价，但承诺“每天低价”就是更大的交易了。正是承诺的不可逆转特性使得战略决策面临着风险。如果战略决策太容易实现，或者改变的成本太低，那么战略决策就没有风险了。格玛沃特的观点和前一章讨论的第一个如果-那么条件直接相关：没有不可逆转性，就无从创造财富，正是不可逆转性使财富的创造充满风险。

标准的战略方法取决于两大基础假设：首先是你能够对未来何种战略会成功作出自信的预测；其次是你可以作出最终形成持续竞争优势的战略性承诺。企业每天都在根据这些假设投资数十亿美元，不幸的是，这两个假设都是错误的。

决定历史走向的小概率事件

有些人热衷于“替代历史”，想象如果偶然的小概率事件出现另一种结果，历史会如何演变。我最喜欢的一个例子是从物理学家默里·盖尔-曼那里听来的。[10]　19世纪末期，“水牛比尔”科迪创造了一出表演，叫做“水牛比尔的蛮荒西部”，主要展示枪法、马术和其他牛仔技能。表演团队中最著名的演员是一位名叫菲比·摩西的女士，她的昵称为安妮·奥克利。据说安妮12岁时就能将一只奔跑的鹌鹑的头部射掉，在“水牛比尔”组织的表演中，她展示了射击术，包括射灭蜡烛火焰和射掉瓶口的木塞。在压轴演出中，安妮会宣布她将射灭一位男士口中所叼的香烟，并邀请观众中勇敢的志愿者上台参与。因为没人有这样的勇气上台，所以安妮会让自己的丈夫弗兰克事先“埋伏”在观众里。他会“自告奋勇”，和安妮一起完成这个把戏。1890年，当蛮荒西部表演来到欧洲时，年轻的王储（后来的皇帝）威廉也在观众之中。压轴表演到来时，令安妮惊讶

的一幕出现了，硬汉王储竟然自愿站了出来。未来的德国王储上到台上，将香烟叼在嘴里，然后站好。安妮前一晚在当地的酒吧中喝到很晚，被这意想不到的情景弄慌了神。她瞄准香烟，扣下扳机……击中了目标。

很多人都设想过，如果当时安妮的手略微地抖一下，那么第一次世界大战可能根本不会发生。如果第一次世界大战没有发生，850万士兵和1　300万平民的生命就会被拯救。如果安妮的手抖了一下，第一次世界大战就不会发生，希特勒就不会从战败的德国废墟上崛起，列宁也不会推翻腐朽的沙俄政府。整个20世纪的历史进程可能会被关键时刻一只手的微微颤抖改变。不过，在那时，没有人会知道这件事的重要性。

安妮·奥克利的故事就是盖尔-曼所称的冻结偶然事件——一个决定历史走向的小概率事件。[11]复杂适应系统的非线性、动态特质意味着，虽然这些事件很小，但历史的走向受之影响可能会很大。

商业历史也建立在诸多冻结偶然事件之上。现代最著名的事例之一发生在1980年的夏天，当时IBM公司正准备发行第一台个人电脑并进入由苹果公司打头的市场。IBM公司接触到了一家位于华盛顿贝尔维尤的只有40名员工的小公司——微软，由哈佛大学辍学生比尔·盖茨和他的朋友保罗·艾伦负责运营。当时IBM想和这家小公司谈谈为新的个人电脑创建一版程序设计语言BASIC。在他们的见面中，IBM让盖茨给出建议，比如新的电脑应该安装什么样的操作系统。盖茨建议IBM去和数字研究室的加里·基尔代尔（Gahy Kildall）聊一聊，后者的CP/M操作系统已经成为微型计算机领域的标杆。然而基尔代尔对来自IBM的人员疑虑重重，当IBM派人来见他时，他去玩热气球了，留下妻子和律师与不知所措的经理见面，并且指示他们连保密协议都不许签。

受挫的IBM经理回来找到盖茨，问他对OS项目感不感兴趣。虽然从未写过OS，盖茨还是给出了肯定的答复。他转身就花了5万美元从一家叫做西雅图电脑产品的公司手中买到了Q-DOS系统，稍加修改后作为PC-DOS卖给了IBM公司。在IBM和微软确定最终协议时，盖茨要求修改一点内容。他想保留将MS-DOS版本的DOS售卖给非IBM计算机的权利。盖茨给IBM提出了一个价格，而且比起软件销售，当时IBM公司对个人电脑的硬件销售更感兴趣，所以双方就同意了。协议于1981年8月12日签署，后来的事情众所周知，他们创造了新的历史。今天，微软公司市值2 700亿美元，而IBM公司市值为1 400亿美元。

这个故事充满了冻结偶然事件：如果基尔代尔没有去玩热气球呢？

如果西雅图电脑产品公司不同意转卖Q-DOS呢？如果盖茨没有在最后一刻更改协议呢？任何一个小的改变都有可能改写商业历史进程。不过，就像安妮·奥克利和她的观众一样，当时没有人能预见那些小事件能带来怎样的历史性结局。

“未来并非昔日模样”

你可能会回答说，不论在经济学中还是在战略计划中，不确定性都是老生常谈。[12]传统经济学的大部分精力都用于研究不确定性，战略计划制定者会运用包括场景分析在内的各种工具来控制不确定性。[13]然而，我们必须记住，传统经济学对于人们面临着何种不确定性这个问题作出了一个大的假设。当经济学家们制作不确定性的模型时，总会把它塑造成随机的样子。[14]这是因为随机性是有迹可循的，你可以利用统计规律对随机体系进行预测。然而，我们在安妮·奥克利和微软公司的故事中看到的不确定性并不是那么有规律的，不太像我们在第8章看到的随机性。随机事件扮演了一个角色——手指的颤抖、见面的机会，是动态的、非线性的系统结构本身赋予了小事件影响历史进程的能力。比起普通的随机性，对这种不确定性而言，更好的描述方式是间断均衡理论和幂律法则。

在复杂适应系统中，间断均衡理论和幂律法则的结合被设计出来哄骗人们进入自我感觉良好的状态，随后却又令人感到不可思议。人类是优秀的模型识别者，我们倾向于寻找过去的模式，并利用这些模式推断未来，以此作出预测。这种方法多半能够成功，如果不能成功，进化历程不可能将我们的大脑塑造成这样。大部分时候，世界是相当稳定的。传统的战略分析常常能够告诉我们很有用的东西，帮助我们作出短期的精准预测。两个标点中间频繁出现的“稳定”时期只能带给我们错误的预测，没有什么比“稳定”更危险。[15]同样，幂律法则的存在意味着，比起人们的构思模型根据过去的经验给予我们的期待，大的变化更容易发生。[16]

接下来的场景曾在一个又一个行业中重复上演，[17]新兴的企业逐渐壮大，成长为大公司。企业高层经理人运用模型识别看到自己的成绩并说道：“啊哈，我们已经了解这个行业了。”然后，企业领导者的模型识别大脑就开始推演未来，并设想出战略计划。战略有了成效，企业就能持续获得成功。行业越是保持相对稳定，管理团队就越是确信自己真的

了解行业情况并能够预测商业走向。管理团队越是自信，就越会精准而明确地调整自己的资产、技能和人力资源，以便在经验充足的情况下取胜。随后有一天，安妮的手颤抖了一下——一种新的技术发展起来，比如竞争对手突然有了一个想法，或者消费者的口味开始转变，而高层管理团队几乎没有人意识到这一点。从这里开始历史出现了分支，世界也开始改变。一开始，企业领导者的模型识别大脑不敢相信变化真的在发生，它们认为正在发生的问题只是一种暂时性的异常，需要更多的证据——毕竟它们已经了解所处的行业很长时间了。然而，当经理团队等待更多证据时，世界一直在继续改变，雪崩加速了。企业现在已经到了一个拐点，管理团队突然发现自己置身于新的游戏，却运用着错误的心理模型、错误的资产和错误的技能。如果变化不是太大或者太快，那么企业就可能幸运地渡过难关；如果变化又大又迅速，企业就会破产，或者被更适应新环境的人接管。

这样我们就需要面对一些问题。一方面，战略计划需要人们对未来做预测，以作出战略承诺。另一方面，未来拥有无限种可能，而我们要走的路可能是由一系列无法预知的冻结偶然事件决定的。与此同时，变化的阶段性使人们的模型识别大脑认为世界比其真实的情况更稳定。更令人沮丧的是，正如我们所见，对抗不确定性只是战争的一部分。

可持续竞争优势的秘密

在17世纪和18世纪，英国东印度公司拥有最大的战略阵地。今天的微软公司和它比起来就像小夫妻店一般。东印度公司完全垄断了四个国家的贸易，拥有全世界趋之若鹜的商品——从咖啡、棉毛织品到鸦片，拥有自己的军队和船队。当公司的商业利益遭到威胁时，皇室会为其发动战争，该公司还统治着世界上1/5的人口。英国东印度公司本该成为当时“最令人羡慕和惧怕”的公司，然而，虽然它有着规模经济和垄断特权，还会残忍压迫当地人，但在19世纪期间，它的竞争优势在技术革命和新的竞争对手到来时土崩瓦解，1873年，这家公司宣告破产。[18]虽然英国东印度公司曾盛极一时，但最终世界改变了，它还是死去了。

所有的竞争优势都是暂时的。[19]虽然一些优势持续的时间较长，但所有的优势资源都有一个期限。这听起来像是老生常谈，但在建构可持续竞争优势并持续超越行业同行，成为“优秀”企业的道路上，人们常常会遗忘这一点。我的书架上塞满了关于这类伟大企业的励志故事书，里

面介绍了这些企业在成功的道路上做了些什么，以及一些企业如何通过做类似的事情走向了成功。其中包括一些经典之作，比如汤姆·彼得斯（Tom Peters）和鲍伯·沃特曼（Bob Waterman）的《追求卓越》（*In Search of Excellence*），柯林斯和杰里·波拉斯的《基业长青》，以及杰克·韦尔奇和拉里·博西迪（Larry Bossidy）这类令人钦佩的管理者的论著。

不过，有趣的是，我的书架的另一侧摆放的都是有关“为何伟大的企业会走向衰落”的书。这部分书包括一些畅销书，比如克莱顿·克里斯滕森的《创新者的窘境》以及加里·哈默尔（Gary Hamel）的《领导革命》（*Leading the Revolution*）。这些书讲述了曾经成功的大企业如何一蹶不振的惨痛案例，以及行业革命、麻烦的官僚主义和骄傲自大如何使得曾经庞大的机构一败涂地。这些书反复重申的主旨是，为了避免重蹈覆辙，企业需要持续创新，并将其作为企业巨头的实力与新兴公司的灵活性结合起来。这些主旨在一些年轻上进的企业中体现得最为明显，它们常常会给注定失败的巨头们致命一击。

有关这些书最有趣的一点是，同样的企业同时出现在两类书中。IBM公司在20世纪80年代初曾是“优秀的榜样”，而到了20世纪90年代它就出现在了“衰落类”书籍中。苹果公司在书架上来回移动，从年轻、锐意进取起步，进入“优秀”区，然后转入“衰落”区，后来在史蒂夫·乔布斯的带领下又回到了“优秀”的一边。令人震惊的是，有相当多“追求卓越”的企业最终彻底从书架上掉落，没能逃脱破产、解体或被收购的命运，比如阿姆达尔（Amdahl）、数字设备公司（Digital Equipment Corporation）、西屋（Westinghouse）、雅达利（Atari）、宝丽来（Polaroid）和凯马特（Kmart）。许多其他的企业，比如施乐公司（Xerox）和洛克希德·马丁（Lockheed Martin）还在活跃着，但是它们在被评为榜样之后表现就一落千丈了。20世纪90年代赢得最多赞誉的两家企业——安然公司（Enron）和世界通信公司（WorldCom）就在2002年以轰动一时的悲剧终结。

也有一些优秀的企业建立了可持续竞争优势，长期保持着高水平的表现。我在麦肯锡公司的前同事迪克·福斯特（Dick Foster）与莎拉·卡普兰（Sarah Kaplan）在他们的《创造性破坏》一书中通过查看最古老的商业排行榜——福布斯100研究了这个问题。[20] 1917年，《福布斯》杂志的创立者博泰·福布斯（Bertie Forbes）发布了第一个“福布斯美国100强企业”排行榜。1987年，为了庆祝福布斯排行榜诞生70周年，他们重

新梳理了之前的榜单，并且调查了原上榜公司的现状。原来榜单上的100家企业中有61家企业以各种各样的方式消失了，不是并入了其他公司，就是破产了。[21]幸存下来的企业中有21家依然活跃，但已经不在100强的榜单上了，只有包含宝洁（Procter & Gamble）、埃克森（Exxon）与花旗银行（Citibank）在内的18家企业依然留在精英团体中。正如福斯特和卡普兰所说，这18家企业是经过大浪淘沙的优秀企业，它们挺过了大萧条、第二次世界大战、70年代的通货膨胀、80年代的并购浪潮以及90年代的技术革命。许多人据此推测它们一定有着极佳的表现，是真正优秀的企业，事实上是这样吗？不是的。除了通用电气公司和柯达公司之外，在那70年中，每一家企业股市价值的增长都低于平均值。1987年以来，柯达公司的表现也下滑了，只有通用电气公司是1918年福布斯100强企业里唯一幸存下来的佼佼者。[22]福斯特和卡普兰还参考了另一个榜单，就是标准普尔500在1957评选出的500家企业，1997年，上榜企业中只有74家依然存在。作为一个整体，这74家企业的标普500指数下降了20%。

在一个长期的框架下考察，以优秀企业为案例的书籍讲述的故事并不是关于可持续竞争优势和持续最佳表现的。相反，它是在通过企业兴衰来讲述竞争优势的转瞬即逝以及市场形势的瞬息万变。

策略是“红桃皇后竞赛”

在2002年和2005年分别发表的两份主要研究成果中，孟菲斯大学的罗伯特·威金斯（Robert Wiggins）和得克萨斯大学的蒂姆·罗弗里（Tim Ruefli）从统计学的角度研究了竞争优势的问题。[23]他们调查了1974年到1997年的40个行业的6 772家企业，根据运营情况将这些企业分为上等、中等和下等的运作者。[24]研究者们更正了许多“优秀”的研究成果中的两个常见错误。第一，威金斯和罗弗里确保了高端企业的优秀表现在统计学上的意义。这就好比许多人在一个大房间里扔硬币，总有人不可避免地一直投中正面，威金斯和罗弗里想确保他们在企业中看到的成功并不只是因为运气。第二，通过运用一系列的滚动5年窗口，他们也确保这种表现是持久的，并不依赖于短短几年的好日子，并且对被检查的特定时间框架不过于敏感。

他们的发现确认了真实的竞争优势既罕有又相对短暂，[25]这些样本中只有5%的企业拥有长达10年或更久的优秀表现。有32家企业，也就

是只有不到0.5%的企业拥有20年的优秀表现，只有3家企业（美国家庭用品公司、礼来公司和3M公司），即0.04%的企业保持着长达50年的优秀表现。威金斯和罗弗里还发现，在这23年的样本选择期里，企业面临的竞争呈加剧态势。竞争优势的平均有效期越来越短，一家家企业在表现排行榜上变换位置的频率越来越快，失去前排位置的概率在研究的时间框架内大约提升了一倍。[26]

威金斯和罗弗里也表示，既安全又稳定的行业是不存在的。他们把样本分成两组：一组为高技术含量，一组为低技术含量，虽然高技术含量这一组的变化速度更快一些，但他们发现同样的模式出现在所有行业中。在描述这一发现时，威金斯和罗弗里指出，熊彼特的灵魂依然存在，“创造性破坏的风暴”比以往任何时候都要猛烈。

威金斯和罗弗里的数据展示出对于以传统经济学为基础的策略概念的巨大挑战，他们所发现的模式在进化生物学家的眼中一点都不陌生。在生物世界，竞争优势也是罕见而短暂的。的确，所有人都认为进化系统理应如此。在生物系统中，物种只能在永无休止的共同进化过程中互相竞赛。正如我们此前所说，掠食动物可能会进化出更快的奔跑速度，而猎物则会进化出更高超的伪装技巧，然后掠食动物又进化出更敏锐的嗅觉，等等，它们都在无限期进化的道路上拼命狂奔。生物学家将这种共同进化的螺旋式上升称为“红桃皇后竞赛”，这个概念得名于刘易斯·卡罗尔的小说《爱丽丝漫游仙境》中的红桃皇后。正如红桃皇后所说：“你必须尽力奔跑，才能保持原地不动。”[27]

想在红桃皇后竞赛中跑赢是不可能的，你能做到的最好的程度就是比竞争者们跑得快。从前文中所讨论的林格伦的进化的“囚徒困境”中可以看出，“什么是最佳策略”这个问题是没有终极答案的。成功的“囚徒困境”策略不可避免会激发其他策略的反应和创新，引起策略生态系统的改变，就这样，引发制胜战略的组成部分会随着时间流逝发生改变。在进化系统中，可持续竞争优势并不存在，只有永无休止的竞赛创造着短暂优势的新资源。

我们对优秀企业的定义已经发生了改变，从长期保持优秀表现到能够长时间将一系列短暂优势串起来。换句话说，就是红桃皇后竞赛中的强势跑者。[28]当然，要在一种优势资源消失之后立刻完美衔接上另一种资源是不可能的，于是我们就会看到这种企业展示出一种模式——先是表现名列前茅，随之被击倒，然后像顽强的拳击手一般回归战斗，再次赢得胜利。威金斯和罗弗里发现了这类企业存在的证据。他们检索了5

年时间里有上等表现的企业的信息库，然后在某刻深入到中等或下等表现的企业，最后再研究上等表现企业在另一个5年内的表现。研究样本包括像强生和默克一样鼎鼎有名的企业，也包括如家庭美元百货公司和伊利诺斯工具公司这样鲜为人知的企业。[29]实现重复优秀表现模式的企业并不多，只占样本数量的1%，但是随着时间的推移，这个比例会不断增长，也许这证明随着竞争越来越激烈，更多的企业学会了如何成为持续的创新者。

创新的不是企业，是市场

持续几十年的优秀表现是世所罕见的（少于0.5%），可重复的优秀表现也是极其稀少的（1%），这就是有关大部分企业的残酷真相。市场是瞬息万变的，但大多数企业都无法做到及时应对。

实际上，据社会学家迈克尔·汉南（Michael Hannan）和管理研究者约翰·弗里曼（John Freeman）所言，企业从本质上来说是迟钝的。20世纪70年代，汉南和弗里曼开始了一系列针对市场“组织生态学”的里程碑式的研究。[30]这些研究显示，虽然在市场层面发生了巨大的经济创新和改变，但个体企业层面所发生的改变要小得多。他们的结论是，经济变革的驱动力是企业的进入与退出，而非个体企业的适应。例如，半导体工业未能从20世纪50年代的晶体管转变成20世纪80年代的超大规模集成电路芯片，就是因为当时如休斯（Hughes）、Transitron和飞歌（Philco）这样的前沿企业适应了。[31]而后来发生改变的原因则是这几家企业被如英特尔（Intel）、日立（Hitachi）和飞利浦（Philips）这样的企业取代了。原先的十大半导体企业里，只有德州仪器（Texas Instruments）和摩托罗拉（Motorola）在大变革中生存了下来。同样，尽管这种转变可能在技术行业中更戏剧化，但有相当多组织生态学家的研究表明，企业数量的动态变化在所有行业都普遍存在。[32]

正是这种进入和退出的动态模式解释了福斯特和卡普兰发现的那些悖论——作为一个整体，标准普尔500长期上榜的企业的表现低于平均水平，正是新企业不断地进入才使得平均线保持在较高水平。

这些以观察和实验为基础的发现并不会令我们感到惊讶。在第三部分，我们讨论过商业如何在市场经济的进化系统中相互作用。从市场的角度来看，每一个企业都是商业计划空间中的一个实验；有些会成功、扩大，而另一些则会失败、消亡。拥有这些商业计划的企业则会随着它

们的命运起起伏伏。这些发现表明，变异、选择和放大商业计划的流程在市场层面比在大多数企业内部效果更好。正如福斯特和卡普兰所说，“市场创造的惊喜和创新比企业要多”。[33]

企业都是大人物统治集团，充斥着人的小缺点和小问题，而市场几乎是纯粹的进化机器。企业所固有的劣势使它永远无法像市场那样完整地拥有商业计划的多样性，它们也无法完美地借鉴真实市场的选择压力，或者拥有资本市场的无限资源，并将其投资到扩大已成功的商业计划中去。从毫无感情的进化法则的角度来看，企业只是进化磨坊里的谷物。

目前，我已经勾勒了一个令人沮丧的画面。战略就是作出将会产生竞争性优势的长期承诺。但是由于冻结偶然事件，我们无法很好地预测未来，而且竞争性优势是罕有的、短暂的、稍纵即逝的。有证据显示，大多数企业在“红桃皇后竞赛”中都是失利者，无法在市场中加快适应的脚步。

我们能做得更好吗？我们能设计出适合度更高的企业吗？尤其是，我们能避开预测的问题发展出更坚定、适合度更高的方法来制定战略吗？[34]

战略就是试验组合

做得更好的关键在于“将进化引入内部”，让变异、选择和放大的车轮在企业内部转动。我们不应把战略视作基于对未来预测的单一计划，而应把它当作试验的组合，也就是诸多随时间推移而进化的相互竞争的商业计划。[35]我们即将看到这种方法的组成元素，在此之前先用一个案例来解释战略试验组合是什么。

在微软公司的故事中，1987年，盖茨与IBM签订合同后过了6年，新兴的个人电脑行业刚刚经历了爆炸性增长。[36]微软公司是增长中借势最多的企业，但是MS-DOS却走到了尽头。顾客开始寻找一种替代运行系统，以便更好地利用新一代机器的图表及其他更强大的功能。S形曲线的转折正在到来，整个行业的走向充满了不确定性。虽然微软公司取得了成功，但是在1987年，和野心勃勃地寻找赚钱途径的数十亿美元的

巨头比起来，它还只是一个市值只有3.46亿美元的小不点。IBM公司当时正在发展强大的多任务处理OS/2系统；AT&T在带领包括开阳电脑公司（Sun Microsystems）和施乐在内的企业创造大受赞赏的Unix操作系统的用户友好型版本；惠普公司和数字设备公司也在推行自己的Unix版本。苹果依然是一个在创新方面领先同行业的竞争者，它推出的高度图形化的麦金塔电脑也卖得很好。

我们可以想象微软公司此时面临的几种选择。选择一：盖茨可以发起一个巨大的“关系公司存亡”的赌博，理想的结果是在竞争对手凭借他们的系统发展到一定规模之前，投资构建一种新的叫作Windows的操作系统，并且尝试将自己的DOS用户转移到新的标准之下。选择二：退出部分操作系统市场，把它让给更大、基础更稳固的竞争对手，而将自己的关注点投向应用，在这方面微软能充分发挥规模小和灵活性的优势。选择三：卖掉公司或与另一个主要的对手联合。虽然选择三会让微软公司失去独立性，但这样的行动有可能打破力量的平衡，从而取代它选择合作的公司。

所有这些选择都意味着对难以改变的行动方向的巨大承诺，并且蕴含着巨大的风险。从保守的观点来看，盖茨应该选择第一种方案，赢得赌注，这样微软就可以继续主导桌面操作系统，并且在接下来的10年中与反垄断监管机构作斗争。但事实并非如此，盖茨和他的团队做了更有趣的事，他们同时进行了六个战略试验。

第一，继续投资MS-DOS。虽然每个人都预测到了操作系统的消亡，但它依然拥有巨大的用户基础。许多用户对换系统都很谨慎，而且每一版DOS系统都在逐步超越上一版。DOS系统依然在继续改变和进化，并且能在一段时间内满足用户所需。

第二，微软把IBM当作真正的威胁。1987年，“蓝色巨人”依然是硬件方面的巨头，想要夺回对操作系统市场的控制，但同时它也知道单枪匹马是有风险的。正如迈克尔·柯里昂（Michael Corleone）在《教父2》中所说：“亲近你的朋友，更要亲近你的敌人。”盖茨和IBM同意将IBM公司的OS/2操作系统项目转成联合运营。

第三，虽然Unix造成的威胁比IBM小，但依然是威胁。微软与包括AT&T在内的许多公司讨论共同运作Unix。讨论保留了微软的选择权，使微软公司加入了这个进程，但是依然对微软的Unix战略持保留意见。这么做的好处是为Unix的支持者们制造额外的不确定性，延缓他们的进步。

第四，除了参与Unix的联盟游戏，微软还收购了个人电脑Unix系统最大卖家的大多数股票，这家公司名为圣克鲁兹运行（Santa Cruz Operation）。这样，即使Unix未能成功，微软也在市场上拥有了自己的产品。

第五，盖茨并未停止投资应用，在资源紧张的情况下，他依然在继续建构这一生意。值得一提的是，微软在苹果麦金塔电脑软件方面确立了自己的位置，超越了苹果本身成为领头供应商。万一苹果利用市场的间断推广自己的操作系统，这也能为微软提供一种保障。

第六，盖茨将主要的投资放在Windows上，而Windows后来成为全世界最好的操作系统。它在DOS的基础上产生，兼容DOS应用，和OS/2、Unix一样是多任务处理系统，并且和麦金塔一样方便使用。最重要的是，它使得个人电脑操作系统的市场控制权牢牢掌握在微软手中。Windows的成功明显是公司最中意的结果。

盖茨所创造的并非大型赌注，而是战略选择的组合。有一种对盖茨所作所为的解释是，他设立了一个高层次的愿景：成为领先的个人电脑操作软件公司，随后他创造了一系列有可能朝着这愿景进化的战略试验组合。

1987年，Windows远没有可能成为赢家。Windows 1.0版本于1985年发行，但是售卖情况并不理想，Windows 2.0版本于1987年发行，饱受技术和延迟问题的困扰。直到1990年发行Windows 3.0版本，微软的未来才与操作系统市场牢牢绑在一起。“安妮的手”也可能朝另一个方向颤动：如果IBM在OS/2的开发方面抢先一步，如果共同开发Unix的公司曾团结一致，如果微软在Windows开发方面受挫，历史很有可能会走向另一个方向。

盖茨没有试图预测未来，而是参照外部市场上演的进化竞争在微软内部创造了众多商业计划。如此，微软就能在未来不断发展。最终，其他的选择都被叫停或者缩小了规模，而Windows操作系统成为公司的重中之重。当时，盖茨还因为这个组合战略备受批评。记者们称，之所以会这样，是因为微软想不出什么战略决策，在困惑中随波逐流，他们不知道盖茨何时才能下定决心。而且，那些在公司内部工作的员工也没有发现自己正在与同事直接竞争。没有证据表明比尔·盖茨在设计这个战略的时候曾研究过进化论或者思考过适合度景观，然而，不论这个方案是如何发展的，结果是创造出了一个能在历史的迂回曲折中坚守的适应性战略。微软一直保留着这个方法，如今，从互联网到企业计算、家庭娱

乐以及移动设备领域，微软都有一系列竞争性试验组合。

从试验组合方法到战略，我们从中可以学到一些基本知识。首先，管理需要为战略创造一个“环境”。创造一个试验组合需要对当前形势有全面的了解以及有一个管理团队共享的愿景。其次，管理需要区分商业计划的过程，才能得出不同的计划组合。再次，组织需要创造一个模仿市场环境的“选择”环境。最后，需要建立流程来扩展成功的商业计划、淘汰不成功的计划。我们将依次讨论每一个关键点。

环境：创造有准备的意识

我曾与一位非常冷淡、讲求实效的高管共事，他声称自己不相信战略计划，认为那不过是一群“高智商的人在胡说八道”。他也很成功，在严酷的市场环境中经营着大杂烩般的产业，试图在数年发展中不断提升利润。有一天，我看到一个战略计划会议的预备材料，而其中这个高管和他的团队准备的分析是整个公司最好的。鉴于他曾声称不相信战略计划，第二天我问他，为何他和他的团队会花费那么多精力去做分析。他的回复是：“我不相信计划，我这么做是为了能有所准备。”当我从这位从不胡说八道的高管引用的路易斯·巴斯德的名言（“机会眷顾有准备的人”）中清醒过来时，我突然明白了他如禅宗一般的智慧。他解释说自己和团队并没有使用传统的战略分析工具预测未来，而是使用为实时决策提供背景资料的工具，使他们能应对所有可能发生的不确定因素。在后来的讨论中，那位高管解释说，战略计划练习是他们的高管团队交流的重要方式。这能使他们拥有一个针对业务的共同参考框架，对事实的共同认识，以及互相沟通的共同语言。[37]

GE金融集团曾经有位高管也跟我说了一些他在计划并购方面得出的理论。他认为战略计划的意义不在于预测未来，而是一种学习练习，使人们能够为不确定的未来做好准备。例如，他发现自己无法预知何时会有重要的并购机会出现。虽然他从来没有一个计划说“我们要购买X公司、Y公司和Z公司”，但如果X公司确定要出售，他的团队就能迅速地坐在谈判桌上出价，出现意外事件的情况非常少。他的团队成员能做到这一点的原因是他们已经展开过充分讨论，对X公司有深入的了解，知道它将如何影响本公司的经济效益。他们对于并购对本公司来说意味着什么已经达成了共识。

这里举出这些案例并不是要你撕掉你的战略书，只是促使你思考传

统战略分析工具的另一种用途。其目的并不是帮你基于对未来的预测找到一个五年计划的“答案”，而是强调“有准备的头脑”。这意味着你需要思考谁的大脑最需要准备，以及如何更好地完成这一任务。

对于大多数企业来说，这种观点上的转变意味着战略计划流程的重塑。大部分流程都集中于创造和作出决定，而不是学习。集中于学习的计划流程具备三个主要属性。[38]

第一，流程的重点应该放在安排深入的讨论以及主要决定人之间的辩论上。典型的计划流程是下属们向拥有决定权的高管们天花乱坠地讲解演示文稿，而人们在这样的会议里一般学不到什么东西。流程的重点应该是创造一个讨论会，让决策者聚在一起集中解决问题或者相互讨论。[39]这样的讨论会规模要小（如果有太多助理人员参与，高管们就无法畅所欲言了），而且时间要充裕——每个业务单元每年要有一天用于首席执行官与高管团队讨论，这比大部分公司一两个小时的团队会议有效得多。

第二，整个流程要以事实和分析为基础。如果只有观点被拿到桌面上说，那么很有可能每个人离开会议室时依然带着他们走进来时的想法。这意味着在进行战略谈话之前要花费几个月的时间精心准备，而且在员工和顾问的帮助下，高管也需要全身心投入。对共同事实基础的一致理解是交流最宝贵的结果。

第三，还应该有其他用来做决定的讨论会。如果战略流程充斥着对预算、目标和资金分配的短期决定，那么学习就不复存在了。做决定的讨论会应该与战略学习过程连接，也应该有所分别。另外，战略讨论会的重点应该是设置环境以告知试验组合的设计和管理。

变异：你的战略树够茂密吗

典型的战略计划流程致力于砍掉战略决策树的树枝，消除选择以及作出选择和承诺。相比之下，进化则强调创造选择，选择越多越好，不论何时都要让战略树尽可能地茂盛。选择是有价值的，[40]战略试验的进化组合可以给予管理团队更多的选择，也就意味着有更多可能发现对的选择，或者，就像我在麦肯锡的同事洛威尔·布莱恩（Lowell Bryan）所说的“用灌了铅的骰子”。[41]这种做法的目标是下许多小的赌注，只有在不确定性变低时再下大赌注。在不确定性之下被迫孤注一掷意味着一家公司被困在茂盛的战略树的对立面。虽然媒体可能会称赞敢下大额赌注

的首席执行官们的勇气和远见，但他们也会无情地嘲笑那些输掉大额赌注的首席执行官。

进化需要孵化力超强的商业计划才能起效。当然，没有任何一家企业的商业计划多样性可以与整个市场相比，但大部分企业都走向了另一个极端，它们对于本公司各项业务都只有一个计划。提起商业计划的多样性，企业往往会和自己作对，因为研究和创新的需求和利用与执行的需求之间存在固有的矛盾。[42]成功而有效的行动需要有所侧重、纪律严明，它们需要清晰的领导和方向。相反，进化战略需要允许人们同时走向不同的方向，并且验证有风险的想法。同样，执行的反馈循环是紧凑而可预见的，你每个季度都能看到结果。相比之下，战略进化的反馈循环很难预见，可能需要花上数年。人们讨厌裁员，而经理人经常试图通过裁掉多余的人来提高效率，然而根据定义，多样性需要有多余的人员和过剩的生产能力。作为重要和有价值的目标，操作效率的驱动力常常形成一些副作用，即降低战略试验的多样性和在企业内部形成相互竞争的业务计划。

除了创造充足的试验组合，相互竞争的商业计划也需要分散在适合度地形上，以便最大化进化搜索过程的效果。如我们此前所见，进化在探索有相互关系的地形方面的有效性原因之一是它混合了跳跃长度。在生物学中，变异和有性重组的结合确保了短距离跳跃和长距离跳跃的混合。当我们思考商业计划地形中的“跳跃距离”时，需要考虑三个方面：风险、相互关系以及时间范围。风险指的是影响战略试验结果的所有不确定性因素，以及承诺的不可逆转程度。相互关系指的是试验距离已具备的经验、技巧和资产有多近或多远。[43]时间范围指试验成功的预计时间。

在微软的案例中，公司的Unix试验和其他试验相比风险较高，周期较长，并且与公司现有的试验关系较小。相比之下，Windows系统本身在这三个方面较为平均，而OS/2可以说是风险最低的试验，因为在技术方面它比开发Windows更省力。可以说，微软公司在整个地形中下了一些比较分散的赌注。试验组合中需要多少个战略试验以及它们在地形中如何分布关乎判断，与商业环境相关，但是这个数字必须多于大多数企业所采用的1。

有个常见的反对战略组合的观点提出，并不是所有的企业都有能力像微软那样同时进行六个战略。事实上，当时微软公司是行业中规模最小、资源最受限的。当时的情况并不是微软提出了六个关于操作系统的

商业计划，IBM提出了12个——总的来说IBM只进行了一个。另外，并不是所有的选择都受到了同等水平的重视。对于微软公司来说，Windows系统的胜利是其更愿意接受的结果，所以这个提议得到了最多的投资。组织拥有的资源确实是有限的，不过这并不会阻碍它们进行选择的组合；阻碍它们的是经理们的想法以及组织的做事方法和文化。

最重要的是，商业计划的多样性和业务的多样性并不是一回事。20世纪60年代，一个比较流行的观点认为，企业应该创造多样性业务组合以降低风险，例如一种业务可能在经济环境好转时运行良好，而另一种业务可能在经济下滑时表现优异。这样的结果是创造出了不相关的业务表现不佳的组合，许多在后来都以破产或重组结束。威金斯和罗弗里的数据以及其他的研究表明，持续聚焦少数业务的企业拥有显著的性能效益。[44]我所提倡的不太一样，我认为在一个业务内部应该有不同的战略组合。例如，微软的试验组合完全是位于操作系统业务内部的。

选择压力：定下愿景

商业计划选择是另一个令许多企业感到困难的领域。企业倾向于将适用于市场的选择压引入内部，但是如一位高管所言，一旦市场的信号进入企业内部，就会被扭曲得“如同哈哈镜”。企业内部有许多市场上根本不存在的选择压，比如“这是首席执行官最热衷的想法，如果退缩了他会很丢脸”，或者“这是我的部门的地盘”。这些来自内部权术、官僚主义和其他因素的曲解常常并不是恶意的结果，它们只是真实世界的人类组织的产物。正如本书前面所讨论的，心理学家发现人们受制于大量的偏差和系统性错误知觉，以及在社会层级中争夺地位的根深蒂固的本能。这些人类的缺点不可避免地影响了组织对市场信号的诠释。

有两种方式可以抵抗哈哈镜效果。第一种相当明显，即改善从市场流入组织的信息的数量、质量和速度。此类信息包括详细的销售和客户数据、市场调研、客户满意度调查、竞争对手情报等。当市场的反馈缺失或过于粗略时，观念和权术就会填充信息真空并扭曲商业计划选择。

第二种方式更复杂一些。首席执行官和高管团队在通过多种途径在组织机构内部创造选择环境方面扮演了至关重要的角色。正式的途径有补偿系统和晋升流程，非正式的途径有首席执行官们在会议中谈论的主题。这些途径共同创造了一套信号，告诉人们什么样的行为和决定会获得奖赏。当一群中层经理人坐在办公室里，辩论一个商业计划的优点

时，这些信号可以帮助他们思考什么类型的计划会得到支持，什么样的计划不会得到支持。重要的是，这种内部选择环境的设计要经过深思熟虑。首席执行官和高管团队成员只能意识到组织中正在进行的很小一部分决定。许多关键的决定不可避免地发生在高层雷达之外，而经过深思熟虑的内部选择环境确保了那些决定既反映了市场现状，也反映了企业的当务之急。

企业愿景是塑造内部选择环境的一个重要工具。大部分企业都有使命、构想、价值和其他类似的鼓励性声明，对于许多企业来说，这类声明只存在于企业总部墙上的彩色海报里，用来给到访者留下深刻印象，但是这些文字很少影响人们的实际行为。只有少数高管团队能够用企业的愿景将市场信号播散在组织内部，塑造决策。

对于良好愿景如何形成以及如何发展和交流，人们的著述颇多。[45]如果想要让愿景在商业计划选择中扮演有效的角色，我认为需要具备四个关键的特点。

第一，愿景必须印刻在组织中许多人的大脑里。如柯林斯和杰里·波拉斯在《基业长青》中所说，有效的愿景要“吸引人——它能散播出去，抓住人心。它是有形的、充满能量的、高度聚焦的，人们看到后立刻就能理解，不需要过多的解释”。[46]相反，大部分企业的愿景没法抓住任何人的心，听起来平淡无味。柯林斯和波拉斯在研究中收集了大量的愿景，并引用了伟人杰克·韦尔奇在20世纪80年代为通用电气立下的愿景：“成为市场上数一数二的公司，拥有小企业的速度和灵活性。”[47]

柯林斯和波拉斯将这个愿景和（曾经表现优秀的）西屋公司的“质量为先、市场领军、技术引领……”进行了比较。韦尔奇的愿景是清晰、有挑战性、有感染力、令人难忘的。但是语言自身无法确保愿景被印刻到人们的大脑中。韦尔奇将数一数二这个想法与赔偿计划、绩效评估与业务单元联系在一起，他着迷地谈论着这个方式。[48]简而言之，他围绕着愿景建构了一个选择环境，塑造了千万人的想法以及无数决定了哪些商业计划应该保留、哪些应该停止的决策。20世纪80年代，通用电气公司计划推出一种新产品，这项产品将帮助其超越对手进入第二大市场，而这个产品本身会吸引大量关注及投入大量资源。同样，对名列第五的业务进行投资的计划即便利润可观也不会得到支持，除非结果能保证将业务推到排名靠前的位置。韦尔奇确保了愿景所创造的选择压可以形成真实的结果——你觉得韦尔奇的高管们有多少在管理排名第三、第四的业务呢？

第二，对于外部世界对组织造成的选择压，愿景必须捕捉一个重要的直觉。当杰克·韦尔奇于1981年宣布自己的愿景时，通用电气公司正濒于倒闭。当时，这家公司正在进行“集团化折价”，也就是说，从塑料、电器到飞机零件的业务的价值加在一起超过了通用电气公司的总价值。各种“入侵者”正虎视眈眈地寻找可以接管和临近解散的公司。韦尔奇当时面临着来自通用电气公司股东的巨大压力，他们希望他能提升通用电气公司的市值，并取消集团化折价。韦尔奇利用“数一数二”诠释外部选择压，用所有人更容易理解的方式将它融入组织，方式极具感染力及可行性。只提出“我们必须消除集团化折价”是无法完成任务的，而通用电气公司人才选拔过程中最受青睐的极具竞争力的经理们在情感上被“数一数二”挑战所打动。

第三，制定愿景时必须找到平衡。要提供的选择压必须足够具体，但由于需要预测未来又不能太具体。例如，成为“关怀员工的企业”这样的愿景就过于平淡和宽泛，没有人知道如何实践这样的目标；成为“英国北部鲸骨束腰领先生产商”这样的愿景也许是在对英国北部鲸骨束腰的未来市场机会作出莫须有的假设。韦尔奇的“数一数二”愿景就达到了适合的平衡，和亨利·福特的“为大众造车”一样。福特的愿景对于提供选择压已经足够明确，例如：“改变商业计划的某个部分会降低车的成本，使它更能被普通人接受吗？”或者：“这种车有没有新的特点能让几乎没有受过机械训练的人感到值得依赖？”而且福特的愿景足够宽泛，在实现目标的方式方面可以产生大量的演绎-探寻。人无法完全逃脱预测，福特的目标假设这种车会拥有广泛的市场，但是在1907年，只有富人才能买得起车，所以“为大众造车”是一个安全的赌注，如果这种车被造出来，一定会大获成功，这也体现出愿景反映了对外部市场有效的选择压的重要直觉。

第四，好的愿景提供了强大的动力，使企业处于不断进化之中，不断尝试新事物，并且支持试验的准则。在进化系统中，适合度地形上的停滞就意味着灭绝——如果停止试验和进化，你就完蛋了。为了适应，企业一定不能安于现状，不能满足于自身的进步，要持续地寻找和探索。正如英特尔公司的前主席安迪·格罗夫（Andy Grove）所说：“只有偏执狂才能生存。”在通用电气公司的例子中，在韦尔奇定下愿景，声称“成为数一数二的公司并不只是一个目标，它是要求”之后，你就知道处于第三第四的公司也会不断尝试新的事物。

不过，企业中并不存在一成不变的愿景。当外部环境的选择压改变

时，提供内部选择压的愿景也必须随之改变。随着时间的推移，愿景要么实现了，如福特公司的例子，要么实用性发生了改变。虽然通用电气公司从来没有实现每一项业务都“数一数二”，但在20世纪80年代末，韦尔奇和他的团队已经修复或者卖掉了大部分表现差的业务部门，集团化折价已经消失，大部分业务都经营良好。于是，韦尔奇开始转而继续追求其他的愿景，比如全球化和在服务业站稳脚跟。

放大：群蜂智慧

当一家公司蔑视逆境，试图发展和选择有前景的商业计划时，它往往很难将这些商业计划扩大。关于大企业没能在眼皮底下成功培育新计划的故事在坊间广泛流传。根据定义，与更大的市场中的人才与资金池相比，企业拥有的这两种资源非常有限。在对人才和资金的内部竞争中，基础丰厚的业务会不可避免地占据上风。较新的、试验性的业务常常缺少资金，由二等团队负责运营，高层管理时间不足。它们也可能会被吸纳进更大、基础更好的业务中被吞噬掉。大部分企业配置人才和资金的过程比更大的市场要慢一些——股票交易人在一毫秒内就能在企业之间转移资金，而在企业内部这可能要花费数月甚至数年。最后，许多企业改变了自己的资源配置决定，用衡量小而新颖的业务的方式来衡量大型、有基础的传统业务。从根本上说，大企业是以开拓传统业务和策略为己任的，它们将人才和资源都集中于此，还会反复衡量。相反，市场则倾向于寻找和支持新的、正在发展的东西。

策略的进化方式强调尽可能长久地保持策略树的繁茂以及选择的开放性。但是不可逆转性——财富创造的第一个乔治斯库-罗根条件，无法永远避免，人们在某个时刻必须作出难以撤回的承诺。企业高管必须作出最终承诺并分配有限的资源，那么我们如何在进化的环境中作出这样的决定呢？

美国第一资本投资国际集团（CapitalOne）是一家正在发展和创新中的信用卡公司，它就在用试验策略在金融服务地形图上寻找新的适合度高峰。当被问到第一资本投资国际集团如何决定扩大哪些试验、取消哪些试验时，一位高管通过描述蜜蜂的行为回答了这个问题。蜜蜂从蜂巢中飞出来后就开始寻找花蜜的来源。当一只蜜蜂成功找到花蜜时，它会回到蜂巢，在其他蜜蜂面前跳一支来回摇摆的舞蹈。它发现的花蜜越多，舞蹈就越活跃；舞蹈越活跃，就有越多蜜蜂飞出蜂巢前去探索花

田。当这片花田枯竭时，蜜蜂们就会回到蜂巢，等待下一个舞者。[49]第一资本投资国际集团的高管说：“我们会观察‘蜜蜂’舞动得有多活跃，如果它们总在跳舞，我们就会将资源向这个机会倾斜。”[50]

扩大试验的前提包括三个重要的因素：知道试验组合中的“花蜜”在哪里；有一些多余的“蜜蜂”涌向诱人的机会；当情况变化时有能力从一片“花田”转移到下一片“花田”。

第一，我们必须知道哪些计划是有前景的，哪些没有。试验组合必须以明确的市场反馈机制为基础。每一个商业计划都必须拥有清晰、深思熟虑的实现途径以尽可能地实时搜寻数据。许多企业犯下的错误就是将同样的性能标准运用在所有的业务和战略试验当中。这个逻辑的动力是讨好金融市场的欲望，例如，“市场以企业层面的利润和资金回报评估我们，所以我们也用同样的方式来评估企业的所有业务”。但事实上这些评估方式更适合处于发展后期的、成熟的业务，对于新的、更具试验性质的企业也许无法给出合适的参考。在这类业务中，阶段式评估方式，比如雇用关键的高管、赢得早期的客户以及达到预算目标可能更加合适。

第二，财务评估方式常常是市场反馈落后的指标。其他更具可操作性的评估方式包括客户满意度调查、装配时间、每平方米销售额、员工流动率以及返工时间，相比于单一的财务评估方式，后者可能会提供一个更完整、更及时的战略试验图景。你需要关于“花蜜”位置的实时反馈，所以每个战略试验都必须有一个定制的平衡计分卡和评估系统来评估它。[51]我在本书的后面会详细讲述这个概念，简而言之，平衡计分卡就是在特定的业务中使价值创造可视化的一套性能标准。平衡计分卡可以为我们提供多维视角，既包括财务指标，也包括非财务指标。

第三，要使进化起作用，系统必须有闲置容量用来做试验。如果企业的所有资源都用于执行传统商业计划，它就无法再发展新的计划。许多企业犯下的错误就是用预算流程将所有的闲置容量从业务中剥离，并将它置于公司的集中控制之下。一般情况下，公司总部会为新的商业计划预留一些资金。这种方式的问题在于，高管的带宽不可避免地局限于改善资源分配决定，他们早已不是对小规模试验成败的最佳判断者了。于是，最终计划就会变得太少太大。将闲置容量推到业务单元之中，并且赋予这些业务单元创造和使用资金支持试验组合的责任与义务是非常重要的。不过，高管团队依然需要作出一些较高层次的分配决定。在崎岖的相互关系地形图中，高峰总是倾向于靠近其他的高峰，所以资源更

多地被分配给了以平衡计分卡标准看来更有成功迹象的倡议，而很少分配给那些处于低谷的倡议。这也可能意味着资源分配模式的转变，因为资源的分配方式从基于目前的业务转变为基于新兴倡议的成果。

这就引出了第四点，资源在试验中的分配需要具备灵活性和可变性。如果无法作出实时调整，那么实时反馈、平衡计分卡和决定权的下放都毫无意义。然而，很少有企业能在一年当中根据反馈来调整资源的分配——每年的预算都是固定的，应急过程一般持续不了多久，耗资也有限。由于集中的企业流程原本就无法快速解决这个问题，要更及时地进行资源分配的唯一方式就是将更多的对资源的自主控制权推向组织各处。商业领袖必须明白，企业中闲置的资源是什么，然后才能评估这些资源的作用，而不是简单地将闲置资源投入“平常的业务”。如果员工无法更好地使用闲置资源，领导们应该将其收回。领导应该将资源从没有收益和前景的试验中撤出，并进行重新配置。

适应心态

企业领导者应该成为勇敢的将军，能预测未来、设计宏大的战略、领导员工参与辉煌的战斗，最后在失败的小型战斗中被解雇。只有勇敢的高管才能反击这样的心态，转而承认未来所固有的不确定性，并且强调在预测和计划中学习与适应。

的确，一个人的心态可能是创造战略适应方案最重要的因素。建构适应方案的其中一种方法是像风险投资者一样思考。[52]从本质上来说，风险投资公司就是战略试验组合。它们的组合包含具备分散风险、关联性以及时间范围的各种投资。风险投资组合在同一个业务部门中常常拥有不止一种投资，本质上是为了尝试各种商业计划并确信至少有一个投资能找到高适合度的山峰。风险投资者的位置介于适合度地形图上广泛的市场范围与大部分狭小的企业范围之间。风险投资者们还会迅速扩大成功的企业并不惜一切代价减少损失。风险投资者倾向于聚焦投资的短期表现，他们还会使用一个广泛、长期的投资命题，并从目标出发提供选择压。很少有风险投资者敢于预测自己投资的产业，他们也不会花费时间制定传统的战略计划。风险投资者们会选择使用组合在摸索中学习，在商业世界风险最高、变化最快的市场中获取高收益。

传统智慧认为，创新心态也意味着容许风险、模棱两可与失控。[53]我所提到的方式并不需要这样的观念，它需要重新定义这些概念。创新

是下许多小赌注，只在起作用的东西上下大赌注，而不是下许多充满风险的大赌注。如果在选择好计划和坏计划中有清晰的认识的话，同时执行多个商业计划并不必然意味着模棱两可。高管们无法控制战略试验的结果，却可以控制创造、选择和扩大试验的过程，以增加成功的概率。

在许多方面，适应心态就是传统愿景的反面。适应心态是高度实用主义的。比起预测明天，它更重视今天的事实，也不期望一切都按计划进行，比起大失败更能接受大量的小失误。总之，适应心态表达的是“我们学到了新东西，要作出改变”。

回望20世纪80年代晚期，当英国石油公司作出一系列关键战略改变时，时任首席执行官布朗（Browne）形容这个过程为“一系列的脚步，一步接着一步，但没有预定的计划。我们做了该做的，当前行或者遇到问题时，自然就明白下一步需要怎么做”。随属公司首席执行官罗德尼·蔡斯（Rodney Chase）补充说道：“这不是一条直线，我们出过错，也不是一切都顺利。但是我们意识到，每一次失误都能使我们获益匪浅。”[54]

16

组织：
创造“心智社会”，释放个体的创造力

西屋公司和通用电气公司是20世纪初美国工业的两大巨头。[1]西屋公司的创立者是乔治·威斯丁豪斯（George Westinghouse），通用电气公司的创立者是托马斯·爱迪生，这两个人都是极具号召力的发明家。西屋公司和通用电气公司都是铁路和发电技术的先驱。这两家公司在大萧条期间损失惨重，在20世纪40年代的战争中为美国战事服务，在战后繁荣时期大量生产冰箱、收音机等商品。20世纪50年代和60年代，这两家公司成为美国家喻户晓的巨头。

20世纪70年代，战后繁荣开始衰退，这两家公司面临的压力日益增长。西屋公司的发展逐渐停滞，利润缩减，沦为一个庞大的、官僚主义盛行的巨兽。在挣扎中，西屋公司开始发展一些不相关的业务，比如房地产。20世纪80年代初期，西屋公司关停部分业务和工厂。到了80年代末，公司抛售了大部分运输和工业生产业务。20世纪90年代中期，西屋公司尝试最后一搏，收购了哥伦比亚广播公司电视网，试图将自己重塑为一个传媒公司。但这一切都显得徒劳无功。到了千禧年，乔治·威斯丁豪斯在任时的荣光不再。哥伦比亚广播公司被维亚康姆集团（Viacom）收购，其余业务被逐一出售，西屋成为英国核燃料有限公司的一个部门。

20世纪70年代和80年代，通用电气公司也经历了寒冬，不得不与成熟的市场、缓慢的增长和官僚主义倾向对抗。不过通用电气公司打赢了这些战斗，并且在20世纪90年代跻身于世界一流公司之列。在这些年中，通用电气公司经历了巨大的改变，从一个工业和耐用品产业集团发展成为后工业金融服务、传媒和技术公司。20世纪末，通用电气成为世界上最有价值的公司。

社会结构与适应性

为什么根本上属于同一个产业、历史相似、处于相同战略位置的两家公司会落入截然不同的命运呢？为什么通用电气公司有能力适应，而西屋公司失败了呢？有些人喜欢将通用电气公司首席执行官杰克·韦尔奇在这期间的超凡领导力视作关键的因素。但是那些仔细研究过通用电气公司的人，包括韦尔奇本人都声称事实远远不止于此。[2]

通用电气公司在其漫长的历史中不断地改造自己，韦尔奇在20世纪80年代对公司的复兴只是其中的一次而已。通用电气公司的历史中充满了失误，比如1893年就曾遭遇现金短缺，同时期西屋公司的交流电技术成为美国标准。通用电气公司曾在大萧条时期遭遇严重的经济困难，20世纪50年代发生道德丑闻，70年代还经历了萎靡时期。但是，正如我们在前一章所定义的优秀模式，这家公司在遇挫之后又数次坚强地站了起来，并且显示出随着经济变化改造自身的非凡能力。

在成立早期，通用电气就已经超越了公司的概念。20世纪初，在首席执行官查尔斯·科芬（Charles Coffin）的领导下，通用电气公司成为一个拥有强大的企业文化和价值的机构。正如吉姆·柯林斯和杰里·波拉斯在《基业长青》中指出，西屋公司从未能够建构这样的机构基础。[3]韦尔奇既是通用电气公司文化的产物，作为首席执行官，韦尔奇也是公司价值最有力的捍卫者。韦尔奇把通用电气公司的员工、结构和文化的结合称为公司的社会结构，并且坚信它是公司独特的运营和适应能力背后的秘密。[4]韦尔奇还自信地提出，公司的社会结构是任何竞争对手都难以模仿的，他不介意公开讨论这个秘密。[5]

本章的主题就是论证韦尔奇的正确性，以及企业的社会结构设计在一个组织的适应性方面扮演的主要角色。我们对社会结构的定义包括三个要素：

1．组织中的个人行为。

2．在追求组织目标时将人与资源匹配起来的结构和流程。

3．组织中的人在与彼此和环境交流时产生的文化。

在探索高效社会结构的构成之前，我们首先需要审视什么是组织，以及为何组织常常抗拒改变。这样我们就可以开始讨论组织是什么，它们为何存在，并且将它们放入进化框架中。随后我们将分析组织在适应方面遇到的多种阻碍，发现一些根深蒂固的阻挠组织改变的因素。最后，我们将看到企业的社会结构设计如何消解那么多阻碍，尤其当社会

结构更少依赖于结构“硬件”，更多依赖于文化“软件”来驱动理想行为时。通过建构有效的社会结构，公司就能将进化带入内部，成为“红桃皇后竞赛”中跑得更快的跑者。

组织——复杂适应系统

如果让一群学者给组织下定义，你会收到一大堆各不相同的答案，经济学家、社会学家、心理学家、人类学家以及法律学者都会给出自己的答案。这些领域都从不同的视角研究组织，每一个视角都有独到的见解。[6]但是有一个视角将这些领域穿插在一起，提供了一个总体的概念：组织是复杂的适应性系统。[7]

组织是由彼此动态互动的一个个实体组成的，随着环境的改变，实体的行为准则和互动网络也随之改变，实体的互动自然产生了行为的宏观模式。当一个人说起“福特这个季度业绩很好”或者“索尼是一家创新公司”时，他就是在谈论成千上万人的行动和互动自然产生的结果。企业是复杂适应系统，嵌套在经济体这个更大的复杂适应系统之中。

组织作为复杂适应系统的分类被证明是十分有用的，但仍需要进一步细化。其他人类群体，例如堵车中的司机，也是一个复杂适应系统，但是我们不会把他们称为组织。要做好这个区分，我们将使用北卡罗来纳大学的社会学家霍华德·阿尔德里奇（Howard Aldrich）的定义：组织是有目标导向、持续边界和社会建构的人类活动系统。[8]

换言之，组织是为了某个目的而设计和建构的。组织的目标提供了推动行为的力量——要缩短现实和理想之间的差距就要采取行动。人类组织中的目标可能是赚钱（商人）、帮助儿童（慈善家），也可能是赢得冠军（运动队）。共同目标和协调活动的存在可以帮助我们区分通常意义上的组织与其他人类群体，比如堵车中的人们或者朋友圈。[9]

阿尔德里奇的定义还指出，组织是具有持续边界的。针对这一点，德国社会哲学家马克斯·韦伯（Max Weber）曾在20世纪早期提出了组织辨别成员和非成员的不同特点，企业会仔细地搜寻潜在的员工，市民会通过选举产生进入政府机构的政客，宗教有各种欢迎新成员的仪式。总之，我们可以将组织视为一个开放的热力学系统，它的内部世界与外部世界有一个明显的区分界限，而且组织的目标驱动着降低组织系统内部与外部环境相关联的熵的活动。

正如阿尔德里奇所说："组织为完成工作具备活动系统，包括原材料、信息或人。"[10]在第三部分我们提出，组织对物质、能量和信息进行了热力学上的不可逆转的改变，将高熵输入转变为低熵输出。这种转变是基于组织目标进行的，因此，我们可以将组织视为创造"健康秩序"或财富的工具。

企业为何存在

在第三部分，我提出企业是经济进化中的互动者。反过来，企业是在一群人控制下的一个或多个业务单元的集合。因此，在我们的定义中，企业是发展和执行商业计划的组织，其共同目标是赚取利润。

那么为什么会存在企业呢？[11]此前，我提到合作具有1＋1＝3的非零和魔法。通过细分任务和合作，组织可以做成单个人做不到的事，并收获成果。这回答了"为什么我们需要组织"的问题，却没能回答"为什么我们要成立组织"。为什么我们不做自由职业者，只合作完成具体的任务，完成之后就散伙呢？为什么我们要创造更稳定的结构，比如合伙企业或公司呢？

这些充满挑衅的问题是1937年由一个年轻的工业律师罗纳德·科斯（Ronald Coase）在一篇开创性的文章《企业的性质》当中提出的。[12]科斯的答案虽然简单但很有见地：人们创造组织是为了降低"交易成本"。如果一个盖房子的人和一个木工只需要合作一次完成一个特定的任务，那么他们最好只签一份合同，然后努力合作完成工作。但是如果他们会重复性地合作很多次，扮演多种角色，那么重复性地为每一次任务和角色起草新的合同代价就很高昂了。在这种情况下，他们有一个更省钱的选择，就是在一个组织中建立一种长久的关系，或是作为合作伙伴，或是一方雇用另一方作为员工。总之，科斯说，如果做自由职业者更省钱，人们就会做自由职业者，但如果将人们集中在一个组织中更省钱，他们也会这么做。

我们可以根据进化框架更新科斯的观点：组织不只是一些情况下更省钱的合作机制，它们还可以使人接触商业计划设计空间的部分，而这是仅凭合同无法获得的。简而言之，组织是探索商业计划空间更好的工具，也是创造财富更好的工具，对此有四个原因。

第一是经济学家所称的"不完备合同"的问题。[13]不论律师有多么聪明，都没有合同可以覆盖所有可能的偶然事件，因为生活总是太复杂。

合作越复杂，合同就越有可能不完备。不完备合同的问题将自由职业者限制在了商业计划空间相对简单的部分。

第二是经济学家所说的抢劫问题。[14]如我们讨论过的，财富的创造需要难以撤销的承诺，这些承诺常常以对特定资产投资的形式出现。和名字暗示的一样，这种资产被设计用来为生产流程做一些特定的事，很少用于其他目的。例如，比萨饼烤箱很适合用来烤比萨，但是烤其他东西就不太实用了。不完备合同的问题意味着自由职业者总是面临一个风险，即合同被解除，有人囤积了昂贵和难以售出的特定资产。其他的自由职业者就可以“劫持”资产的主人，威胁退出，除非他们能在交易中获得更多利益。通过创造对资产的共享所有权，并将参与者的利益关联起来，组织就减少了投资在特定资产中以及作出难以撤销的承诺的风险。

第三是组织为合作提供了一个结构，即便在特定代理改变的情况下合作也能长期存续。有些项目需要很多年才能完成，例如发现、试验以及生产一种新的药品可能需要10年甚至更长的时间。而人类个体又会制造不便，比如离职、跳槽、生病甚至死亡。如果每次有人离职，葛兰素史克公司（Glaxo Smith Kline）都不得不解散或者重新开始，那么它就可能一无所成。通过创造允许个人代理参与的合作结构，组织使我们触及了更复杂的商业计划。

第四是组织为集体学习提供了工具。我们习惯于将学习看作个体活动。然而，法国科学家皮埃尔·格拉斯（Pierre Grasse）在20世纪50年代进行了一列试验后声称，组织也可以学习。[15]由于人类组织有些复杂，格拉斯在一个更简单的组织形式——白蚁蚁群中进行了试验。白蚁在筑巢时，会先建造柱状的支撑结构并促进通风。令人震惊的是，这些柱子之间的距离均等，高度也相同。假设白蚁不具备太多个体智慧，沟通能力也有限，格拉斯很好奇白蚁是如何学会这么做的。通过仔细观察，格拉斯确信白蚁是通过共同完成作品来学习的。当一只白蚁开始建设一根柱子时，它就在其他白蚁身上启动了“如果-那么”的指令，它们就会参与进来。一根柱子的存在又会激活其他的指令，阻止白蚁在特定的区域建造柱子，所以这些柱子之间的距离均等。从本质上来说，白蚁“读取”了植入柱子的信息，并作出了相应的回应。例如，如果一根柱子比较矮，白蚁就会添加更多的土；如果一根柱子比较高，白蚁就会去建造下一根柱子；白蚁会避免在已有柱子的一定的半径范围内建造其他的柱子。格拉斯将这种形式的学习定义为非直接通信。[16]

人类也会做相同的事。我们会合作创造包含植入信息的人造物，并

且根据这些人造物中植入的信息来改变自己的行为。例如，一群建筑师正在设计一幢建筑，其中一个人可能修改了蓝图，并把它传给另一位同事做进一步的修改。这位同事再把蓝图交给工程师，工程师又将它拿给了其他人，等等。设计在一种重复性群体学习的过程中被修改完善。蓝图包含了植入的信息，是集体学习的工具。组织中充满了文件、图表和其他人造物，使主体不断变化下的非直接通信学习成为可能。组织在为计划提供仓库或者说集体记忆方面扮演了关键的角色。

学习和适应之间是有区别的，学习是在追求目标的过程中获得知识，而适应是在回应环境带来的选择压时作出改变。虽然适应需要获得知识，学习也可以发生在对环境压力的回应中，但我们会发现区分这两个术语还是有用的。我们将看到，虽然组织通常善于学习，它们在适应中却容易遇到更多的问题。

基于科斯的看法，我们能看到比起由契约自由职业者组成的经济形式，由组织组成的经济形式具有巨大的优势。组织可以使我们到达商业计划设计空间的广泛范围。随着组织方式的进化，我们就能创造越来越复杂的组织，它们反过来又使我们可以发现和执行越来越复杂、创造的财富越来越多的商业计划。[17]

执行与适应

任何经济组织都面临两个基本的挑战：第一，它必须执行现有的商业计划去应对今天的挑战；第二，它必须适应这些计划来应对明天的挑战。

执行与适应是进化系统里任何设计的终极命令。[18]在热力学第二定律控制的世界中，在当前环境中成功地找到合适的位置是生存下去的必要条件——摄入的卡路里必须多于消耗的卡路里，赚的钱必须多于花的钱。不过，据我们所知，进化系统中战略的保存期限可能很短，所以在环境变化时必须持续地探索新的战略，或者冒险将自己置身于窘境。但是由于执行和适应的过程都需要资源，系统内部自然存在压力。要投资多少资源用于今天的执行，又该投资多少资源用于适应明天呢？在商业组织中，短期内运营的需求和长期投资与创新同样关键的需求之间存在对钱、人和高管时间的持续竞争。

1982年，汤姆·彼得斯与鲍伯·沃特曼在《追求卓越》一书中关注了这个挑战的管理学意义，他们是率先注意到这个问题的著名管理学作

家。[19]他们认为，组织必须对执行“紧”，对适应“松”。这种辩证法一直以来都是管理学写作的中心议题：柯林斯与波拉斯在《基业长青》中描述了“控制”与“创造”的重要性，迪克·福斯特与莎拉·卡普兰在《创造性破坏》中提出了平衡“运营”与“创新”的需求，迈克尔·塔什曼（Michael Tushman）与查尔斯·奥雷利（Charles O’Reilly）在《创新制胜》（*Winning Through Innovation*）一书中描绘了他们眼中的“二元性”组织。[20]在学术界，有关这个话题最著名、引用最多的论文是1991年由斯坦福大学的詹姆斯·马奇创作的，他使用了两个著名的术语——“探索”与“开发”。[21]由于每位作者的用语有所差别，这个对立的阴阳主题也就不断出现。不过，正如我们在上一章所见，有证据表明企业普遍更擅长执行，而非适应。

有些人可能会争辩说，这只是事物的自然法则——经济进化在市场层面能更好地操控，组织只应该是执行机器。我们应该让市场去探索，让企业去开发。但是我们必须知道，企业内部的适应性越强，社会就受益越多。适应减少了企业失败和员工失业所带来的损失。当一家企业适应失败时，不妥协对个体生活和当地社区都会产生消极的影响。另外，组织适应能力的匮乏并不只是企业需要关注的问题。许多其他种类的组织，包括非营利组织、政府机构和军队团体，都面临着同样的问题。[22]

现在，我们对于组织是什么以及组织为何存在已经有了了解，接下来来看一下为什么组织在执行与适应之间总是倾向于执行。鉴于组织是复杂适应系统，我们将从三个层面审视它。

1．个体主体层面。我们的心理模型为何使得个体难以适应？

2．组织结构层面。层级、复杂的灾难以及资源的配置是如何为适应设置障碍的？

3．自然发生层面。企业文化如何在自然发生的组织行为中扮演关键的角色？

这三个层面回应了我们定义社会结构时用到的三个元素：行为、结构和文化。实质上，社会结构的目的就是处理这三个层面，打破适应中的障碍，并且使组织在短期运营和长期进化之间达到平衡。

个体：透过玫瑰色的玻璃看世界

一个组织的适应能力某种程度上反映了组织中的个体心理适应的功能。这个问题并非只属于位于顶端的人，还属于层级之中的人——如果只有首席执行官“感受到”改变的需求，其他人都毫无感觉，或者只有年轻的管理者看到了这个需求，而执行委员会的那些老人们却意识不到，那这种感受就是毫无益处的。感受到改变的需求也是充满挑战的。近年来认知科学的研究表明，人倾向于在同一个心理模式下学习。虽然人类擅长在现有的心理模式上增加新的信息，他们却难以在基础的层面改变这些模式。简而言之，我们倾向于画地为牢。

我们来看看个体适应能力低的四个原因：第一，人们倾向于过分乐观，这会减弱对改变需求的感受。第二，人们还倾向于规避损失，使得他们不愿意冒特定的风险。第三，我们对世界的归类以及建构心理模型的方式会阻碍改变。第四，改变的间断均衡论性质更青睐于僵化的领导风格，而非灵活的领导风格。[23]

要改变心理模型，人必须首先觉察到改变的需求。他们必须看到目标，看到当前状况和世界轨迹之间的差距。另外，基础性的改变常常需要人们相信一些不好的事情即将发生，只有改变才能避免。“9·11”事件发生之前的几个月里，美国政府内部对事件的描述告诉了我们之前范例中的心理模型：中央情报局的分析家还处于冷战模式；联邦调查局的特工更擅长对抗犯罪头子，而非恐怖分子；政治家们通过民族国家的棱镜来看世界。[24]尽管基地组织的威胁已经暴露无遗，但直到惨剧发生，人们才在震惊中开始作出必要的适应。人们通常倾向于积极地描绘现实，忽略掉令人不适的事实，只有意外的挫折才能使他们看到现实并不是一切都好。

澳大利亚管理研究生院的研究者丹·洛瓦洛（Dan Lovallo）和普林斯顿大学的诺贝尔奖得主丹尼尔·卡尼曼将这种面对消极事实的乐观称为“妄想乐观”。[25]他们指出了管理预测和计划中存在的持久的过度乐观。例如，兰德公司的一个针对44个化学加工厂建设项目的研究发现，平均而言，加工厂的耗资是预期的两倍，而产出只有预期生产能力的3/4。在真实世界中，预测常常失误，项目延期、技术效果不如预期、并购没有希望中的成功。过度乐观的倾向是人类内心深处的欲望之一，我们渴望相信自己的能力。洛瓦洛与卡尼曼引用了另一项研究，其中100万学生参与了调查，70%的人认为自己的领导能力超过了平均水

平，只有2%的人认为自己的领导能力低于平均水平。

造成妄想乐观的原因之一是管理者是因乐观而被选中的。乐观主义者往往也是更易相处的同事，而悲观主义者可能会被认为是不忠的。正如洛瓦洛与卡尼曼所说："传递坏消息的人更容易被其他同事孤立、躲避和忽视。当悲观的观点被压制，乐观的观点被奖赏，组织的批判性思考能力就会被削弱。"个体员工的乐观倾向会互相加强，对未来的不切实际的乐观会因此被群体确认。[26]

这就意味着距离组织阶梯的顶层越近，我们就会发现越高比例的乐观主义者，这就为适应制造了障碍，因为乐观主义者会比现实主义者更少感受到改变的强烈需求。

洛瓦洛和卡尼曼推荐的解决方式是确保外界的观点和数据能定期被输入组织。例如，如果你想建造一个化工厂，就需要看一看其他44个例子的记录。我们可以通过培养现实主义文化来对抗过度乐观主义。杰克·韦尔奇就曾因为在自己的员工中戳破过度乐观的泡沫而令人生厌，他的其中一个观点就是管理者必须"面对现实，看到事实的本来面目，而不是他们所希望的样子"。[27]

个体：适应与规避损失

丹尼尔·卡尼曼还论证了人有另一个妨碍适应的特点——规避损失的倾向。例如，如果你能通过抛硬币决定自己是赢1 500美金还是输1 000美金，传统经济理性告诉你应该放手一搏，但大量的实验表明，大部分人不会这么做。人们倾向于认为赢1 500美金是"锦上添花"，而输掉1 000美金就是"灾难临头"。对损失给予更多的关注在非金融场合也同样适用，比如有些人在做重要演讲时，更关注的是自己是否看起来像个傻瓜，而不是担心自己所讲述的内容。

规避损失的普遍倾向使得公司在探索时更加谨慎，更关注开发。探索从定义上来说是更加不确定的活动，带有更多的负面风险。还有一个更加微妙的效果是，在判断风险时，比起绝对规模，人们更倾向于关注相对规模。一位首席执行官可能会将一个1亿美元的决定视作大赌注，而将100万美元的决定视作小菜一碟，并且倾向于不去规避损失。不过，一位中层管理者看待100万美元的决定时可能会和首席执行官看待1亿美元的决定时一样担忧。这就意味着两种决定都会被以相同的规避损失层级对待。不过从逻辑上说，组织对于1亿美元的赌注更应该规避损

失。但是对损失100万美元的中层管理者的惩罚（比如解雇）对其个人而言与输掉1亿美元赌注的首席执行官所承受的痛苦一样消极——实际上，由于绑在金色降落伞里，首席执行官通常都能更稳健地着陆。正如卡尼曼所指出的，这种下赌注方面的相对主义意味着组织常常对下层所投的小额赌注过于谨慎，而对于高层所做的大额赌注则不够重视。

针对小额赌注的规避损失常常出现在试验需求面前。在前一章我们看到，当进化拥有许多小额赌注来试验不同层级的风险时，进化的效果最好。当公司拒绝小风险组合时，它们就会在战略上陷入困境，最终将所有的筹码都压在大赌注投资和并购之上。减少中低层管理者对小额赌注所负的责任有利于开展更多的试验，发展出更高的适应能力。

个体：经验的代价

在第6章，我们介绍了认知科学家约翰·霍兰德和他的同事如何将人类心理模型描绘成了“默认层级”。[28]概念被组合在层级结构中（例如，小的、蓝色的、嗡嗡叫的东西被归类为“苍蝇”），当我们拥有特定信息时，我们的心理模型就会启动相应法则，而当事情不确定时，我们就会回归默认法则。

默认层级拥有一个巨大的益处，它允许人们在拥有不完整信息时采取行动。而且，它允许人对自己的行动进行微调，以及在获得更多特定信息和经验时进行学习。从不利的一面看，这种结构会使经验成为一把双刃剑。[29]在年轻、经验不足时，我们的默认层级是比较肤浅的，对世界的看法也更单一。这种思考方式既有优点也有缺点，优点是心理模型很容易改变。我们总是在准备吸收新的经历，也很容易辨认出层级，因为本身要辨认的东西也不多。缺点是我们对设定的情况作出正确回应的可能性更低。于是，社会中对于年轻人头脑适应力更强，但更容易作出不适应的事情的印象并非毫无根据。

随着年龄的增长，我们会获得更多经验，默认层级日益增加，情况就倒过来了。我们的特定经历增多，得到许多关于什么有效、什么无效的反馈，心理模型发展出了高度复杂的种类、互通法则和权值结构，我们变得不再容易将所有事物都视为新事物，反而试图将新知识与旧的经验联系起来，并且将这些经历合并进已有的经验中。当偶尔遇到经验领域之外的事物时，我们不得不创造出一个新的、高层的经验分类或者重新组合已有的分类。随着心理模型变得越来越复杂，这样的大变动就越

来越困难。重组一个更老、更有经验的心理模型就像重组通用汽车公司，而重组一个更年轻、经验不那么丰富的模型就像重组一家新的小公司。随着时间的推移，我们的心理模型越来越稳定，需要越来越大的力量将它们撼动。

我们在一生之中都在平衡探索与开发。在形成自己的心理模型并且从环境中获得反馈时，重点在于探索，但最终我们会形成一种确定的规范，重点就转为开发已经发展出的模型。这个论断并非“年龄歧视”，当然也有20岁的人守旧古板，也有适应力强的70岁老人。但是从广义上来说，我们的心理模型结构是随着时间的推移而改变的，在每一个发展阶段都存在优势和弱点。

在大部分时候，稳定的心理模型在稳定的环境中都如鱼得水。然而，如此前所述，在复杂适应系统中改变的间断特性是有害的设计，它会哄骗我们的模型认知、建立法则的头脑获得稳定感，然后用巨大的改变袭击我们。企业倾向于采用层级化的组织方式，使最有经验的人处于最高的位置。这种安排展现出一种缺陷：处于顶端的心理模型更擅长在稳定的环境中运行，却在探索方面能力不够，对环境变化的适应力较弱。结果可能就是环境变化发生时在组织中形成巨大的惯性。这种心理模型惯性解释了为何情况转变会导致高层管理者大量的变动，换人比转变人的心理模型简单多了。

如果经验心理模型中对改变的抗拒是人类认知的深层特点，我们该如何防备它呢？一个好的起点是位于顶端的人认识到它。经营组织的人必须意识到，虽然他们的智慧可能具有很高的价值，他们对改变和灵活性的敏感性却可能很低。而且，他们的资历可能会吓到资历浅的人，这些人通常对环境中的变化更敏感，并且持有不同的观点。也许针对这种情况最重要的解决方式就是在高层管理团队和董事会中安排不同年龄和资历的人。执行委员会和董事会里全是白发老人可不是好兆头，同样，如果没有白发的人也不是好兆头。但年龄并不是唯一需要具备多样化的因素，经历也是一样——不同的产业、功能背景、国际经历、创业经历、公司经历、非商业经验，等等。高层多样化的心理模型增加了一种可能性：至少有一个心理模型能捕捉环境中的重大变化，应对具有创新性的想法。不过，要使多种心理模型有效协作，高层领导团队必须接受一种非层级文化，欢迎挑战和辩论。一位独裁的首席执行官很容易抵制一个本应充满生气、多元化的管理团队。

个体："僵硬"VS."灵活"

1980年英国保守党的年会上，首相玛格丽特·撒切尔全力维护了政府，这是她职业生涯中的著名演讲之一。在经过了充斥着罢工和暴力抗议的"寒冬"之后，撒切尔大胆地拒绝了偏离自身想法的经济改革路线，她声称"对那些正屏气凝神等待最喜爱的'U型反转'一词的人，我只有一句话要说：你想反转随你便，但铁娘子我绝不会转"。[30]撒切尔是一位传奇的领袖。对于英国应该如何被领导，铁娘子有着清晰的愿景和意识形态，她很少偏离自己的路线，对批评的声音更是充耳不闻。有人因她意志强大的领袖精神而爱戴她，也有人因她的不敏感和毫不妥协而憎恶她。

与之相反，美国前总统比尔·克林顿则常常被媒体称作"伟大的华夫饼"（Great Waffler）。他常因没有固定的意识形态、愿意妥协、放宽选择以及对民意调查十分在意而备受批评。然而，正是这样对环境敏感的品质使得克林顿在艰难的第一任任期之后又被选为阿肯色州州长，并且于1994年中期选举失利后在1996年赢得了第二个总统任期。这些品质还使他在明显分化的国会中推进了一系列备受争议的项目，从福利改革到北美自由贸易协定。一些人爱克林顿的倾听技巧、适应力以及建立共识的能力，另一些人则无法忍受他的搪塞和玩弄手段。

两种典型的领导——铁娘子与克林顿，严厉的发布命令者与灵活的共识建立者，都有各自的优点与缺点，都有各自的伟大成果与巨大失败。这些也是我们自身经历中可能遇到的领导者类型，考虑到严厉与灵活领导风格的存在，我们需要探究的是这些相反的风格是否和严厉与灵活的组织有关。

这是约翰·霍普金斯大学的经济学家约瑟夫·哈林顿（Joseph Harrington）在自己的研究中提出的问题。他首先建立了一个相对简单的模型，描绘了层级化的组织。主体进入层级的底层，基于表现被晋升或解雇，而进入顶层的人最终会在强制退休年龄离开。[31]哈林顿的模型中有两种主体，严格的玛格丽特·撒切尔型和灵活的比尔·克林顿型，层级中的主体必须在每个阶段作出一个决定，选择战略A或战略B。对于外部的环境，A或B都是更好的选择，每个主体的表现部分由他的决定与环境的适合程度决定。

严厉的主体天生就偏向于某一决定，他们常常使用同一种策略，一直坚持A或一直坚持B，不论外界环境如何变化。相反，灵活的主体会考

虑环境的需求，根据情况变化选择A或B。

你可能以为灵活的主体每次都会赢，因为他们的决定总是会适应环境的需求。但实际上还有另一个因素在起作用——经验。也就是说，如果一个主体在他一生中曾使用过5次战略A，另一个主体只使用过3次，那么更有经验的主体会做得更好。总体上主体的表现源自使策略适应环境以及经验两方面的作用。哈林顿随即描述了一个决定，即在层级中设置一个特定的主体，比它表现好的主体会被提拔到顶层，比它表现差的主体会被解雇。你认为谁会赢得胜利呢？是严厉者还是灵活者呢？

答案是，这取决于环境如何。如果环境十分随意，在A和B之间来回跳跃、毫无偏好，那么灵活的主体就会赢，并且有可能占据组织的整个上层部分。在这样的环境中，顽固的主体在每一回合都有一半的概率出错，很少能等到积累足够的经验。

然而，如果环境稳定，并且对其中一个战略有所偏好，比如偏向战略B，那么严厉的主体就会赢。原因是严厉的主体在这个例子中使用战略B有高于一般的胜算，比每一回的成败机会均等的主体表现更好。另外，每当严厉者使用偏好的战略（假设是战略B），遇到使用相同战略的灵活者，严厉者会因其具备更丰富的经验获胜——严厉者一直使用战略B，而灵活者只是偶尔使用战略B，即使环境需要战略A，有时严厉者也会存活，因为它每个回合有1/3的概率遇到另一个使用战略B的严厉者，在这种情况下，赢家要么是具有更多经验的、使用战略B的严厉者，要么是平局，要通过抛硬币定胜负。于是，组织的上层就更可能由使用偏好战略、想法一致的严厉者组成。在稳定的环境中，经验很重要——不是所有的严厉者都会进入顶层，但他们往往更容易在层级中上升。

这是个非常简单的模型，但它捕捉到了真实世界中组织的一个动态实质。如果环境偏向于某种战略，频繁使用这种战略的人就更容易升到顶层，打败适应力更强的人。但是，当环境改变时，组织就会充满停滞不前的人。

哈林顿在自己的模型中加入了一剂现实主义。他抛弃了不断在A和B之间变动的环境，转而使用更像我们所讨论过的改变的间断均衡模式的环境。在间断均衡模式中，战略A和战略B会成群出现，这样环境就可能在一段时间内保持稳定（例如都是A），然后经过一段时间的反复无常（例如在战略A和战略B之间游移），后来又在一段时间内保持为新的模式（例如都是B）。哈林顿发现，间断均衡模式进一步强化了严厉者对

顶层的占据。直观的解释是在长期、稳定的阶段，使用适当战略的严厉者因其丰富的经验拥有一项重要优势，于是他们升到了顶层。在拐点到来之前，一切都十分顺利。然后，之前的严厉者突然无法使用正确的战略，被淘汰出局。在转折过程中，一些灵活者暂时掌控局面，但当环境最终稳定，新一代的严厉者升到顶层，循环重新开始。尽管哈林顿的模型十分简单，但这个模式和许多大公司经历的管理模式改变如出一辙。

哈林顿指出，转变越不频繁、越突然、越严重，严厉者就越能占据主导，当环境开始变化时，情况就越糟糕。在持续的、反复无常的环境中，灵活者做得更好。哈林顿衡量了转变过程中的模型组织的经济实力，发现经济损失的量和组织在稳定阶段试图抓住的灵活者数量有着直接的关系。赶走灵活者，被严厉者占据的组织在稳定阶段可能会表现很好，例如实现盈利目标，但是当拐点到来时，一切都变得糟糕起来。相反，储备了一部分灵活者并且将两种类型的人更好地组合在一起的组织在稳定时期表现不佳、效率更低，但是在转变期却能更平稳过渡。

结构：多少个层级

我们现在将更进一步探讨从个体主体转向组织结构对适应力造成的影响。杰克·韦尔奇曾称这种结构为组织的“硬件”——汇报关系、管理流程以及企业资源配置的设计。[32]在本书前部分，我们讨论过层级在组织中扮演的有益角色。它使网状结构以一种便捷的形式组织起来，作决定不再需要每个人都与其他所有人沟通，因此避免了复杂的僵局。层级过少的组织会难以改变和适应，因为沟通的密度太高，达成共识的需要减慢了作出决定的速度。第7章中有一个标题叫“为层级稍做欢呼”，而不是热烈欢呼。在这个部分，我们将看到，层级在某些时候也会减弱适应力。

斯科特·佩奇（Scott　Page）是政治学家，也是密歇根大学复杂系统研究中心的副主任，他研究的问题是：为何组织会进化出不同数量的层级？[33]为何有些组织是扁平化的，有些则恰恰相反？佩奇假设组织结构与组织解决的问题的性质有关。组织需要解决五花八门的问题，从如何制造一种产品或提供某种服务到如何招聘新人和如何管理公司财务。这些问题的困难程度也大不相同。例如，协调两个人尽力将屋子打扫干净可能相对简单，而协调几千人去建造空客A380就比较复杂了。

佩奇提议，经济协调问题的难度可以从两个角度进行分析。一个是

它是在解构问题，以便同时解决多个困难。例如，房屋清洁工可能会分工合作，一个人清扫厨房，另一个人清洁浴室。平行任务的困难程度由完成最艰巨的工作所需的时间来衡量，例如打扫厨房可能需要更长时间。第二个是需要连续操作的步骤的数量。例如，在组建A380时，工程师需要仔细地确定油箱与控制面板的组装顺序来制造机翼，然后确定整架飞机的组装顺序。所需的步骤越多，找到达到目标所需的正确顺序就越难，耗时就越长。问题中的每一项任务都需要很长时间，又有大量的任务需要谨慎地排序，就算得上一个困难的问题了。

佩奇假定，组织会进化以便匹配它们试图解决的问题的性质和难度。粗略地说，如果问题可以很容易地解构成平行的任务，并且不需要太多排序，那么组织就会倾向于扁平化。如果问题不能轻易解构，并且具有数量众多的步骤，那么组织层级就倾向于窄深型。例如，在律师事务所中，小团队的律师可以同时代理多个客户的案子，团队之间的协调相对较少，所以律所的组织机构相对来说是平的。实际上，大部分大型律所在下级同事与上级合伙人之间只有四到五层。相反，设计和建造一架喷气式飞机则需要谨慎排序，因此，需要许多协调者和管理者以确保组装成功。生产空中客车的公司这样的组织在前线铆接工人与高管之间往往具有许多层级。

层级与困难程度之间的这种关系不仅解释了为何不同行业的组织可能具有完全不同的结构，还向我们展示了组织内部的执行和适应可能需要不同的结构。概括来说就是执行任务包括协调众多复杂的有序和平行的流程，因此需要深层层级。相反，探索活动通常需要扁平组织。战略试验组合通常由一系列小团队组成，在总体组合层面有一些基本的协调，但团队多半以平行和相对自主的方式工作。执行和探索问题的不同性质制造出了天然的矛盾。长于执行的组织具有庞大的深层级，用以解决巨大而复杂的问题，擅长控制、效率和追责。它们并不是被设计用于同时承担许多小任务的，因为这些小任务的重点在于速度、灵活和自主，对于冗余和失败有较高的容忍度。

我们很少看到探索活动从主流的执行层级中有机地产生，这反映了组织结构和问题性质之间的不协调。这也是为何在许多公司中，当发现探索需求时，高管们就会从层级中分出小团队，将它们放在不同的大楼里，并创造出各种臭鼬工厂以追求创新倡议。这种倡议可能非常成功，例如苹果公司的麦金塔电脑就是由一个小型团队自主研发的，这个著名的团队在自己的大楼桅杆上挂上了海盗旗。

从执行层级中分离出来的臭鼬工厂可能是解决失谐的唯一途径，但是这个途径容易掉入两个陷阱。第一，如果所有的探索倡议都由集团中心来操控和赞助，组织会不可避免地只探索适合度地形的有限部分，即高管雷达屏幕上的部分。第二，总有一些问题将这些倡议重新组合进主流组织。当然，如果成功的倡议最终需要扩大，那么重组就是必要的。如果除了集团组合，业务单元也拥有资源和自主权自己进行试验组合，那这两个问题就都比较简单了。这使得公司能更广泛地探索适合度地形，并为成功倡议重组进入组织提供了更自然的途径。

结构：资源和商业计划的共同进化

早在探索与开发之间的矛盾在管理圈爆发之前的1959年，伦敦经济学院的一位经济学家伊迪丝·彭罗斯（Edith Penrose）就出版了一本短小但影响力巨大的书——《企业成长理论》（*The Theory of the Growth of the Firm*）。[34]彭罗斯将企业成长视作寻找和探索的过程，管理团队在环境中寻找新机会，然后利用企业资源来开发这些机会。彭罗斯在书中讨论了企业的物质（例如厂房和设备）和人力（例如管理才能、员工技能）资源，而现代理论家们将她的定义扩展到包括无形但同等重要的资源，例如知识、品牌、名誉和关系。总而言之，管理团队会利用资源来开发机会。[35]

彭罗斯的理论有两层含义。第一，企业的资源由管理团队想开发的特定机会决定。如果管理团队在纳米技术中看到了好机会，就会雇用有纳米技术经验的研究人员，购置制造纳米产品的机器，以及创建自己在纳米技术领域的专业品牌和声誉。第二，企业在特定时间所拥有的资源定义和限制了它的开发能力。例如，如果管理团队正在运营一座鱼类加工厂，而首席执行官某天早晨醒来突然沉迷于纳米技术，那么该企业在开发纳米技术方面就会遇到困难。纳米技术领域也许真的有机会，但是企业的资源（擅长切鱼的工人、装罐机，以及"海味"一类的品牌）决定了企业的机会存在于与鱼类加工相关的事情上，而非纳米技术领域。

我们可以用在第三部分提到的术语中重新阐释彭罗斯的想法。[36]管理团队运用演绎-探寻来寻找有利可图的商业计划，但搜索仅限于他们认为自己能执行的计划，而这些又由组织的资源现状决定。但是当计划被执行，管理者会改变资源的配置，也因此改变了可能的未来商业计划的空间。这是一个商业计划和企业资源之间持续不断的共同进化圈。

计划和资源的共同进化在组织结构中创造了路径依赖，同时形成了适应的另一重要障碍。一家鱼类加工公司无法轻松地一跃成为纳米技术领域的新星，即使市场对纳米技术公司的奖赏远多于鱼类加工公司。更准确地说，这样的改变不应该发生。如果鱼类加工公司有现金，投资银行家们也很乐意支持它的纳米技术生意，但是鱼类加工管理团队却没有必需的技能和经验来成功地运营纳米技术业务。西屋公司曾经在从工业公司转型为传媒公司时作出过“驴唇不对马嘴”的尝试，结果是显而易见的。

工业和战略中重大的改变可能会成功，但也是一步一个脚印，而非一蹴而就的。在1990年的经典论文《企业核心竞争力》（*The Core Competence of the Corporation*）中，普拉哈拉德（C. K. Prahalad）与哈默尔描述了佳能公司是如何基于摄影和平版印刷业务的资源转型做影印机业务的。[37]这个行动反过来在激光打印机和传真领域创造了机会。这些机会需要公司具备微电子学方面的能力，这反过来又使佳能成为数码摄影领域的佼佼者，同时产出相机和打印机。随着时间的推移，资源和机会可以互相成就。

资源的路径依赖以及商业计划和资源之间的共同进化圈成为另一个抑制适应力的因素。企业可能很容易就积累了与前进目标不匹配的资源，而重新配置这些资源又要占用更多可利用的时间和金钱。

彭罗斯的理论进一步强调了战略试验组合的必要性。这些组合不仅在组织中创造了更多样的商业计划，它们中的倡议还创造了更多样的资源。启动一项倡议通常需要雇用更多员工、开发新技能、购置资产等。回到微软的例子，为了探索Windows和Unix系统的潜力，盖茨需要分别雇用这两个领域的人才，包括Unix程序员，他还购买了个人电脑已有的Unix系统运行的权限。如果市场偏向Unix，他不仅能从商业计划试验中获得反馈，还拥有了一整套Unix的核心资源基础。没有这项试验，微软公司可能会缺乏Unix系统的核心专家或资产。

文化：行为法则

我把文化放在框架的顶层是因为它是组织的自然特征。正如此前所述，当我们说艾默生电气公司（Emerson Electric）有绩效导向的文化时，我们就是在陈述成千上万个体的决定和行为的自然结果。[38]

文化是一个模糊的词语，历年来围绕它产生过许多定义。[39]在这

里，我将使用它的一个适用于进化框架的定义，与人类学家罗伯特·博伊德和彼得·理查森的定义相似：

> 文化是一群主体的自然特征，由主体在社会环境中的行动以及互动的行为法则（或准则）决定。文化是在社会中传播和习得的。[40]

简要地说，文化准则就是在社会环境中行动的经验法则。准则是对“应该”的陈述，是关于在特定情况下被社会或组织视作正确、合适的事或被期待完成的事。[41]

我们可以将组织环境中的文化视作一系列同心圆，从广泛遵循的准则到更具体和个体化的准则。最外层的圆代表人类共用的文化准则。人类学家和认知科学家们的研究表明，人类大脑中有一套相对较小系列准则，而这可能是生物进化的产物。[42]例如，谋杀是普遍的禁忌，而所有社会对互惠行为都有奖赏。这些基本准则的特定变化是通过模仿和教授的过程完成的，也就是这样在文化中传播。例如，每个社会对谋杀的定义都是不同的，而我们在第6章也看到，公正的定义在每个社会也不同。[43]

在由外向内的下一个圆中，我们可以看到大量的特定文化准则。这些准则在特定的社会中进化出来，并随着时间的推移在人与人之间传播，由父母传给子女、老师传给学生、老板传给员工以及在好友之间相传，等等。特定文化准则的传播渠道有很多，包括故事、音乐、宗教、文字和媒体。这些准则常常包含各种各样的行为，从性到待客之道，再到如何尊重他人等。当来自不同文化背景的人相互碰撞，这种行为法则就会产生摩擦，而当游客在陌生国家观光时又可以体验到无穷的乐趣。在社会准则的范围内，常常有一些子群和次子群，它们由地理、宗教或其他因素决定，所以我们可以谈论德国文化，也可以谈论巴伐利亚文化或者天主教巴伐利亚文化。

再往里看，我们就能找到组织文化。组织文化常常存在于更广阔的社会文化环境中，更概括地说是人类文化中。包括公司在内的组织会随时间推移发展出自己独特的文化，我们可以将其视作特定组织内的互动法则。例如，一个组织的文化可能允许新人自由发表意见，另一个组织的文化就可能比较专制，要求新人只能服从于上级。或者，在有的公司文化中，承诺是灵活的，不会被严肃对待，而在有的公司文化中，承诺是必须履行的，若未能履行就会产生严重的后果。关键是这些文化是组

织特有的，并非普遍的社会特点。例如，索尼与松下电器都具有一些相同的日本特有的文化特点。然而，在这些共有的特点外，你会发现每个公司都有其特殊的行为法则，我们可以认为这个子集就是组织文化。

许多关于公司和管理的书籍在讨论企业文化时会用到价值这个词，而我们应该区分价值和准则。价值是关于相信什么是最重要的论断，比如"我们认为在团队合作中，平等是有价值的"。准则的论断中有"应该"的成分，例如"你在团队中工作时应该采用非层级化的方式"。价值和准则与信条和行为常常成对出现。许多公司会交替使用这两个术语，但我会用准则或文化法则来强调将其转化为个体行为的重要性。

位于最里面的一个圆代表个体准则。每个人都有一套自己信奉的行为法则。松下电器的员工可能都具备共同人类准则、日本准则以及各种各样的松下组织准则，但他们都是个体，每个人的行为都有所差别。要区别特殊准则和组织准则、社会准则或人类准则，就要明确准则适用的范围。例如，一个人年轻时可能认识到个体倡议是被鼓励的，而另一个人可能相信随波逐流才是正解。如果个体倡议在一个社会中是常见的特征，那么我们可以说它是广泛的社会准则；如果它只是偶尔出现在社会中，但在特定的组织中较为普遍，那么它就是一个组织准则；如果这种行为法则只是偶尔出现在组织中，但特定的个体却具备，那它就是一个个体准则。

文化：十条戒律

从某种意义上说，文化对组织表现的影响是显而易见的。如果所有或大部分公司员工行为都以某一种特定方式呈现，那么这种行为就会影响公司的整体表现。更有趣的问题是，哪些行为是令人满意的，哪些不是？

促使公司取得成功的准则一般是这个公司所特有的，比如它所处的行业、它的特定商业计划、竞争环境等。但是，正如托尔斯泰所说，"幸福的家庭都是相似的，不幸的家庭各有各的不幸"。在对表现好和适应力强的公司文化的研究中，我们发现一些规律性的特点，在这里列出10条。[44]

这并不是一个标准清单，你一定会争辩清单上的某些条目是否应该存在。其实这份清单只是为了说明众多研究者确定的准则类型。我将这些准则分成了三类，第一类与个体表现相关："我在独处时应该如何表

现？”第二类体现了合作行为：“在与他人交流时，我们应该作何表现？”第三类体现了促进探索和创新的行为。

表现准则

1．表现导向。竭尽全力、孜孜以求、积极主动、不断进步。

2．诚实。对他人诚实，对自己诚实，坦坦荡荡，面对现实。

3．精英主义。以价值为基础奖励个体。

合作准则

4．相互信任。信任同事，相信他们能完成任务。

5．互惠原则。己所不欲，勿施于人。

6．共同目标。将组织的利益放在个人利益之上，与所有人同舟共济。

创新准则

7．非层级。新人可以挑战上级，重要的是想法有无价值，而非出自何人之口。

8．开放。保持好奇，独辟蹊径，愿意试验，精益求精。

9．以事实为依据。找到事实，最重要的是事实，而非意见。

10．挑战。时刻准备竞争，竞争无止境。

这些准则读起来十全十美，很容易获得认同。难的是在大型组织的关系网中贯彻执行，让成百上千、甚至成千上万的来自不同背景、有着不同个性的人遵守。

虽然存在重重困难，但也有一些企业在某些准则方面做到了“言出必行”。强生公司的“信条”十分有名，它由罗伯特·伍德·强生（Robert W.Johnson）于1943年写下，从此成为公司的行动指南。这个信条提出了一个共有的更高目标——“服务病人、医生和护士”（准则6），体现

精英主义原则（准则3）和行为伦理原则（准则2）。强生公司的创立者罗伯特·伍德·强生曾声明："失败是我们最重要的成果"，这表明他乐意试验并承担风险（准则8）。[45]

我曾任职的麦肯锡公司也以强大的公司文化著称。执掌该企业59年的马文·鲍尔（Marvin　Bower）于20世纪40年代到50年代创立的一系列准则一直以来都存在于企业新的合伙人的头脑中。[46]虽然词语发生了改变，但核心概念却被保留了下来。麦肯锡主张全球"一家公司"，这个概念表明了互信和互惠的准则（准则4和准则5）。它还"重视精英"（准则3），是一家"非层级化的公司"（准则7）。"对客户产生持续、积极的影响"这一准则赋予了麦肯锡的合作伙伴以共同目标（准则6）。"不进则退"，不晋升就离职的用人之道则塑造了激烈竞争的绩效文化（准则1和准则10）。

通用电气公司也以竞争、绩效导向著称，公司由"A"选手，"修、关、卖"以及领导4E（能量、修整、增能和执行）这样的概念来驱动。[47]杰克·韦尔奇的一句名言就是"无界组织"，直接指向准则8，他的另一句话"面对现实"指向的是准则2，肯定"A"想法的概念反映了准则7。

几乎没有公司可以完全贯彻这十条准则。即便是备受推崇的通用电气公司有时候也无法达成一些合作准则。组织会不可避免地对某些准则进行特定的阐述以及强调不同的方面。关键并不是我所列出的特定准则被刻上石碑，而是我们会发现文化法则如果广泛地在人群中传播，就会对宏观层面产生积极的影响。

文化：内在张力

每一条准则都在执行与适应的故事以及减少组织对结构与流程的依赖中扮演了关键的角色。个体表现准则明显对执行具有重要影响，它们在适应中也扮演着重要角色。如果个体表现准则被深深地植入一个组织，并且成功地驱动个体行为，那么层级与流程则无须过于紧凑和严格就能达成良好的表现。如果层级和集中控制流程不那么紧凑，反过来就使组织拥有了更多用于试验的闲散资源，将责任下放到前线。当个体缺乏准则时，层级就会下压，流程就会变得紧凑，这也许会成功地推进执行，却会损失组织的适应力。

合作准则也对适应和执行具有重要影响。在低信任度、低合作倾向的环境中，人与人之间的互动需要事无巨细的表述，需要遵循无数规

则，灵活度较低。在这样的组织中，员工和管理者关系极差，工会需要制定详细的合同，甚至会规定哪个人应该用哪把螺丝刀。而形成信任和合作的准则可以促使人们开动脑筋根据情况作出最好的决策，既创造了更好的业绩，也增加了试验和改进的可能。

个体表现、信任和合作行为的准则对于适应来说都是必要非充分条件，必须有另外的准则促使个体探索并承担责任。另外，必须有准则保护组织中的个体（例如准则7和准则9），并给予他们创新的空间。

将三种类别中的每一类准则集中，它们就能帮助组织减少“硬件”方面的阻碍，用文化“软件”来代替层级和过程。如果广泛采用和实践，这种准则毫无疑问会对大部分组织产生积极的影响。

在现实生活中，大部分组织都没有或者很少践行这些准则。虽然价值和文化可能被写在公司大厅的公告牌或者职员的桌子上，但是在许多公司它们都不是管理者的头等大事。在贯彻这些准则的公司中，文化会通过员工之间的互动有机地进化，很少有高管介入。这样的组织文化一般都混合了各种优点和缺点，例如一些组织可能在个体表现准则方面比较强，但在合作准则方面较弱，或者相反。十条准则之间存在着内在张力，例如强烈的、绩效导向的精英主义和团队合作之间的张力。除非这些张力在高层领导之间被积极地管理，并且被组织的结构强化，准则的一些子集不可避免地会产生影响。实际上，没有强大的领导力，在个体表现、合作和创新之间平衡的一系列准则就无法有机地产生。

创造适应性社会结构

对于为何组织一般无法像市场一样快速或成功地适应，我们已经找到了一些原因。

个体

- 人类心理模型偏向乐观，限制了我们发现改变需求的能力。
- 人们天生偏爱规避损失，这使得他们在试验中投资过少。
- 由于我们的心理模型会在稳定的环境中获得经验，所以它们会在这种环境中更有效，但代价是灵活度会降低。

- 在稳定的环境中，相对于那些更适合探索和适应的人（灵活者），组织层级更偏爱晋升那些具有执行目前商业模式的技能与经验的人。

结构

- 执行复杂生产和服务流程的挑战促使组织发展出了深入而密集的层级。这些结构无法很好地适应探索的任务，因为探索需要更扁平、自主的组织结构。
- 商业计划与资源之间共同进化的关系限制了一家企业可以探索的商业计划的空间，资源的路径依赖限制了一个组织转向新的商业计划所需的能力和速度。

文化

- 强大的文化准则对于适应力来说是至关重要的，因为它们可以使组织在无须牺牲执行效果的情况下减弱集中式分层控制。
- 文化准则中存在内在张力，能鼓励个体表现、合作与创新，但这些张力必须由组织的高层领导者积极地管理。

执行与适应看起来是不可调和的一对矛盾，我们在前文中讨论过的基于试验的数据表明，大部分公司更喜欢执行而非适应。通过了解这种分裂的缘由，我们也能给出潜在的解决方案。通过将管理的中心杠杆从组织结构的硬件转变为文化软件，我们就能一一解决这些问题。

粗略地说，经营一家大型组织有两种方式。第一是使用管理层级结构。我们可以定义角色、目标、任务以及步骤，然后通过目标的完成情况来衡量个体和组织的表现，同时进行相应的奖惩。这种结构会降低个体的自由度，将他们的行为限制在组织期望的范围内。这种方式的好处在于可控、可靠与可预料。使用这种结构杠杆可以促使人们做出更多被期望做的事。虽然我在形容层级运用时用了一些不好听的字眼，但它在某些方面还是令人满意的，比如使人们拥有了定义明确的角色、有趣的任务以及能够清楚衡量的成功，尤其是在回报丰厚的情况下。没有这种层级结构，人类现在可能还在山洞里钻木取火。

第二种方式依然需要层级架构，但是在骨架上较少使用命令控制肌

肉。组织可以更加依赖于文化而非结构与流程来引导个体行为。文化驱动的组织的首要优势就是文化法则比结构更加灵活。文化法则提供了一般的指引，把“如何完成”这个问题留给了个体。简而言之，它们需要个体开动脑筋。

文化法则可以用来解决我们发现的问题。鼓励人们面对现实、作出基于事实的决定以及经过深思熟虑承担风险这样的文化法则可以抵消过度乐观和规避风险。同样，非层级化的文化可以使新人将说出想法认定为一种义务（麦肯锡称之为“提出异议的义务”），而不仅仅是期望，以此解决高层心理模型的僵化问题。拥抱不同管理风格，同时重视执行与创新技能的文化可以减轻顽固者与灵活者的问题。文化的改变还可以解决青睐执行层级超过试验组合的结构问题。尤其，推进个体表现和合作法则可以使层级结构更松弛，同时避免短期表现的损失，这反过来使得试验的资源和责任被推广到组织内部。

使用倾向于文化引导的管理方式并不意味着要抛弃结构和流程的硬件。相反，正如我们所注意到的，社会结构对两者都需要。它更意味着组织的硬件和软件必须并存并且相互加强。我们可以回过头来问问，如果目标是更好地平衡执行与适应，那么什么样的文化会支持这样的目标？什么样的结构和流程会支持这样的文化？如果一个被期望达成的准则是个人问责，但是对达到目标的人没有奖励，对没能达标的人也没有惩罚，那么准则就只是一句空话。同样，创新与风险承担方面的准则需要预算、人力资源和绩效流程的大力支持。

也许社会结构中最关键的硬件杠杆就是公司的人力资源与培训流程。与和人类的天性作斗争相比，拥有强大文化的公司发现，雇用在文化方面适应的人比在人身上创造适应的文化容易得多。这些公司对于什么样的候选人拥有适合组织的DNA有着清晰的认识，他们会在候选人身上积极地寻找这些特质。不过，尽管新人可能具备适合的素质，但拥有强大文化的公司不会假设这些新人会轻松地通过渗透学习公司的文化。相反，这些公司会小心翼翼地设计介绍的流程，将准则介绍给新人，并且在这个人选职业生涯的培训和其他项目中强化这些准则。通过明确地奖励践行准则的人，不奖励甚至惩罚没能践行准则的人，评估和晋升流程进一步强化了文化。从第一次面试到退休的这些流程中，人们在持续不断地接收公司文化中的信息。

高管的行为和沟通在建设和保持文化引导的组织结构中起到了巨大的作用。这是老生常谈了，但也是一个事实，如果高管不以身作则，践

行准则，那么其他人也不会这么做。如前所述，文化准则的平衡系统并非有组织地产生的，所以所有拥有强大文化的公司在历史上总有一个时刻，当时的首席执行官扮演了文化总设计师和加强员的角色。这些个体也为公司仔细地设计了一系列期望准则，并保证这些准则被高管团队践行，将准则深深地植入组织，然后创造出保证准则在组织中渗透的机制。随着时间的推移，一旦文化被很好地建立，文化设计与强化就能（而且应该）成为高管团队的集体责任，但是在文化建立与形成的关键时期，首席执行官的个人领导力是无可替代的。

交流也很重要，认知科学告诉我们，大部分公司的变革计划都是落后的。[48]人类是固执的生物，不会因为老板的言辞、充满鼓舞性话语的演示文稿或标识牌就改变自己的心理模型和行为。公司变革计划需要处理人类认知中喜爱故事、模式识别、情感化以及善于归纳的方面，而非诉诸基于事实、演绎的方面。人们需要在已有的心理模型中受到震撼，看到现实和需求之间亟待弥合的鸿沟。变革计划需要有一个基于事实的论点，但是交流的重点必须放在故事、类推以及模式上，以帮助人们看到问题。大部分变革计划都十分被动，许多交流都是自上而下的。但学习是交互式的，所以变革计划需要让人们亲自克服问题。

许多高管发现话语片段这个想法有点不合口味，将它和政客联系在一起。但是关于人类认知的事实是，我们会记得使用了押韵、比喻、讽刺或者其他技巧的朗朗上口、有节奏感的短语。杰克·韦尔奇是使用话语片段的大师，他用话语片段将行为准则融入通用电气公司的文化中，取得了非常好的效果。[49]此外，频繁地重复也是关键，这就是为何宗教领袖要每周会见信众，为何政治家们要满脸通红地进行相同主题的政治演说——我们的大脑倾向于通过质量及频率来衡量信息的重要性。

这种管理方式并不轻松，所以也不常见。正如罗伯特·威金斯和蒂姆·罗弗里指出的，拥有适应和重复成功模式的公司只占所有公司数量的1/100。通用电气、惠普、IBM 、强生、英国石油公司、杜邦、宝洁以及高盛是经常被谈起的少数建设适应性社会结构的公司。[50]这些公司都不完美，但都拥有持续性的优良表现，它们从逆境中重生，并且在充满变数的环境中存在了很长时间。

心智社会

伊迪丝·彭罗斯在她的《企业成长理论》中提到一些问题：是否存在

限制企业规模和成长的因素？如果存在的话，包括哪些？[51]她发现了两个因素，其中第一个是管理复杂性的能力。在某一时刻，组织可能会陷入彭罗斯所说的“太大而无法有效管理”的情况。[52]然而，这种限制因素的表现一直在进化，如同我们的物理技术与社会技术的进化一样。很明显，用电脑、电话、飞机和现代管理技术要比用信件、信鸽、轮船和19世纪的标准工时更容易经营一个全球化组织。有证据表明，随着时间的推移，这个限制变松了，因为财富500强公司的平均员工数量几乎增至3倍，从1955年的约16 000人变为2003年的将近48 000人。[53]

彭罗斯发现的第二个限制因素是知识。一家企业的知识面有多广，它就能增长多快。[54]知识是进化过程的产物，我们可以说，企业增长的速度受限于进化过程。虽然经济中的组织规模增长了，企业规模和增长率的分布却相对稳定。根据布鲁金斯学会的罗伯特·埃克斯特尔和波士顿大学的吉恩·斯坦利的研究，两者的分布都符合幂定律。[55]这些幂定律的存在表明，系统中的确存在限制因素。你可以假设，令我们停滞不前的也许正是彭罗斯所说的第二个限制因素。我们的物理技术和社会技术使得组织发展出了协调、控制和执行的能力，远远超过了它们的创新和创造知识的能力。

如果这是事实，那么就意味着我们的操作远未发掘人类的潜力。伟大的认知学家和人工智能研究专家马文·明斯基说，我们所说的“智能”并不是一个单独的东西，而是从多个个体的集体互动中产生的突生现象。智能的魔力在于当这些部分以特定的方式组织时，它们能作出个体无法完成的事。明斯基称对智能的这种描述可以概括为“心智社会”。[56]

人类组织经历了三个发展阶段。第一个阶段是社会技术的进化，它使得陌生人可以在信任的基础上展开合作。第二个阶段是大规模组织的创造，它能开发催生产业革命的物理技术。也许我们现在正处于第三阶段的初始阶段，刚开始学习如何在大型组织中创造“心智社会”。

大部分组织只利用了它们的人员脑力的一小部分。工业时代产生的层级组织使人类实现了曾经无法想象的大规模合作，并执行无比复杂的商业计划。不过，这些结构内在的僵化最终限制了它们在市场中的进化。通过文化杠杆，公司可以在组织内部创造真正的社会，解放员工的思想。在任何社会中，自由都意味着责任，文化必须支持表现、执行、探索及适应。创造这样的文化是真正的“社会工程”，其中伴随着与这个短语相关的所有风险和不确定性。对于成功的公司来说，最终的奖赏是它们创造出的制度能够在未来持续驱动财富的创造。

17

金融：
预期的生态系统

约翰·梅里韦瑟（John Meriwether）是掌管数十亿美元对冲基金的长期资本管理公司（Long-Term Capital）的首席执行官。1998年夏天，他在休假期间接到了一个紧急电话。8月份的世界金融市场犹如不停向前行驶的过山车，亚洲经济体陆续崩塌，美国市场出现大幅波动。到月底，这列过山车彻底脱轨。长期资本管理公司在几天内损失了44亿美元，[1]梅里韦瑟满含羞耻地被召回康涅狄格州的格林威治（Greenwich），开始实施紧急救助计划。

诺贝尔奖溃败

人们会记住1998年的夏天，这一时间节点标志着旧的金融理论时代的结束，新的理论时代的开始。1987年股市暴跌撼动了象牙塔的根基，1998年的一系列事件使得股市彻底崩盘。市场的波动令人难以置信，由此及彼、接连崩溃，人们对于风险的认知突然发生变化，传统经济学始料未及，几乎无法解释此类现象。这一切也极具讽刺意味：长期资本管理公司的投资策略的主要设计者也是两位现代金融理论之父——罗伯特·默顿（Robert Merton）和迈伦·斯科尔斯（Myron Scholes），他们也是诺贝尔经济学奖获得者。正如长期资本管理公司溃败后的第二天默顿指出的，“根据我们的模型推测，不可能出现这种情况”，然而它确实发生了，而且还刊登在了《华尔街日报》的头版。[2]

新近发生的种种事件只会引发人们对传统经济学中的金融理论更多的质疑。传统经济学整齐、理性、均衡的世界中并不存在股市泡沫。然而，在1997年至2000年的牛市期间，美国金融市场中的美元价值增加了超过12.7万亿美元，然后在世纪末的巨大崩盘中，10.8万亿美元迅速蒸发。人们再也无法忽略现实世界和教科书之间的差距了。

经济学领域中只有极少数学科的理论和方程式直接来自学术界并应

用于现实世界，而金融学正是其中之一。投资者、银行、企业经理和政府决策者都在大量使用传统金融工具。以金融学为基础进行决策的包括数十亿美元的交易、两家公司是否合并以及中央银行如何设定利率。因此，任何对传统金融理论错误的指控都有着重要影响力。

金融理论独一无二的原因是它是经济学领域进行实证分析最多的理论。金融经济学家拥有丰富的资源——每分每秒进行的成千上万宗资产交易数据。他们也很幸运，因为金融市场已经运行了很久，并保持着良好的记录，经济学家不仅可以随时查看数据，还可以查看数十年的历史数据。不幸的是，所有这些数据都不符合传统理论。

被遗忘的法国人和积满灰尘的图书馆

1900年，法国索邦大学（Sorbonne）一位年轻的数学系研究生路易·巴舍利耶（Louis Bachelier）完成了一篇题为《投机理论》（*The Theory of Speculation*）的论文。[3]他在论文中提出了股票价格随机游走理论。正如本书前面所描述的那样，随机游走是指某个事物在每个时间增量上以随机方向移动随机距离。巴舍利耶的理论根本的含义是，随着时间的推移，股票价格的变动路径和醉汉在巴黎街头游荡一样，并没有什么规律可循。不幸的是，巴舍利耶并没有说服论文审查委员会的老师们，接下来的60年里，他的理论没有引起任何注意。[4]

随后，1954年，一位名叫吉米·萨维奇（Jimmie Savage）的统计学家在芝加哥大学图书馆里查找资料，刚好发现了路易·巴舍利耶在1914年写的一本小书。萨维奇被这本积满灰尘的书深深吸引，于是给颇具传奇色彩的经济学家保罗·萨缪尔森写了一封信。萨缪尔森从索邦大学获得了巴舍利耶的论文副本，在萨缪尔森的影响下，随机游走理论成为传统金融理论的基石。

在萨缪尔森重提巴舍利耶的理论之后的30年里，包括萨缪尔森、保罗·库特纳（Paul Cootner）、哈里·马科维茨（Harry Markowitz）、詹姆斯·托宾（James Tobin）、弗兰科·莫迪利安尼（Franco Modigliani）、默顿·米勒（Merton Miller）、费希尔·布莱克（Fisher Black）、迈伦·斯科尔斯、尤金·法马（Eugene Fama）、威廉·夏普（William Sharpe）和罗伯特·默顿在内的一众经济学家进一步发展了传统金融学的主要观点。1973年，普林斯顿大学一位名为伯顿·马尔基尔（Burton Malkiel）的教授出版了《漫步华尔街》（*A Random Walk*

Down Wall Street）一书，这本书完整地总结了这一时代的研究，并向华尔街专业人士和个人投资者推广了这些研究。[5]这本书后来成为畅销书和工商管理学硕士课程的固定读物，整整一代交易者和投资者在被遗忘了的法国人提出的观点下逐渐成长起来。

教科书上挑选股票的方法

根据传统金融理论，对股票或其他任何金融工具进行估值时，应该看投资者未来从股票中获得的现金流。[6]假设一家名为可预测公司的公司需要支付股票的股息，而我们有一个水晶球，可以完美地预测该公司在未来5年内每年支付10美元的股息。水晶球还告诉我们，在第6年，诺贝尔公司将以每股100美元的价格接管可预测公司。通过持有该公司的一部分股票，诺贝尔公司5年间每年将收到10美元的股息，相当于50美元，加上第6年的100美元，总计150美元（不考虑税收）。因此，可预测公司未来每股的价值是150美元。然而，众所周知，当今一美元的价值比未来一美元的价值要高。如果只是为了在6年内收回150美元，人们当前就不会以150美元的价格购买该股票。人们希望钱能够获得回报，而回报取决于投资的风险以及替代方案如何。

现在想象一下，假如你没有水晶球，只是期待可预测公司能够按照刚刚概述的方式运转。你已经阅读了公司的所有报告，分析了公司在市场中的地位，并且利用所有的公开信息来作出决定。但是，虽然你希望可预测公司以某种方式帮你赚更多，但你也知道它可能变得更糟，也可能变得更好。同时，你也可以将资金放到拥有政府担保，几乎没有风险的计息银行账户中，而不是投放到面临相关风险的可预测公司中。因此，假设你认为可预测公司的风险太大，需要获得10%的回报才值得出手。如果是这种情况，那么当前你只愿意支付94美元，以便一段时间后可以收回150美元，这样的投资可以让你获得10%的回报。[7]当投资者认为当前美元价值超过未来美元的价值时，这种想法被称为贴现（discounting），贴现金额取决于投资者认为其正在承担的风险水平。

在传统金融理论中，所有投资者都是理性的，并按照理性行事。他们会针对特定股票搜集所有可用的信息，形成一套对通过持有投资以期在未来获得现金流的预期，评估风险以得出贴现率，然后计算当前股票的价值。因此，在给定时刻，股票的价格反映了股票中所有的可用信息，以及每个人基于该信息的预期。假设你预计可预测公司会增长，其

股息每年会从10美元上涨至11美元、12美元、13美元或14美元。那么当前你就会愿意支付101美元，而非94美元，来获得10%的回报。因此，预计某支股票会上涨的投资者其实已经将这种预期纳入了当前的价格之中。同样，预计未来股息将下降也会导致估值下降。同样，对投资风险的看法发生变化也会引起贴现率发生变化，投资人对股票赋予的价值也会随之发生改变。

一旦掌握了所有可用的信息并计算出愿意支付的价格，从理论上讲，你就不会改变出手价格了，除非有新的信息出现，改变你对未来的预期。因此，如果你听说可预测公司另一个股利支付为10美元时，这正符合你的预期，你会继续在该股票上投入94美元。但是如果情况相反，你得到消息称可预测公司正在增加股息，这让你始料未及，那么你的估值就会有一定幅度的上涨。

到目前为止，我们已经描述了传统金融理论中个人投资者作出决定的过程，接下来要考虑当若干投资者进行交易时会发生什么。每个人都会以不同的方式来解释信息。乐观的投资者也许会预期可预测公司的估值为101美元，而悲观的投资者可能会将其估值为87美元。乐观主义者愿意以最高101美元的价格购买股票，而悲观主义者愿意以最低87美元的价格出售股票。因此，两者就有可能进行交易。现在，市场上很多人都打着自己的算盘，也持有不同的股票，因此我们可以举行拍卖，让所有人进行交易，实质上也是自行分类，让悲观主义者向乐观主义者出售股票。在某些时候，所有的投资者都会以他们想要的价格持有想要的股票，这时他们将不再有任何交易的动机。在这一节点上，每个人都很满意，没有人想再进行交易，市场处于均衡状态。拍卖师此时设定的价格就是市场价格。实际上，市场价格代表市场中所有的预期和信息达成的共识。

在这个故事中，市场应该只是以市场价格处于均衡状态，直到新的信息进入为止。我们可以想象每个人都在敲手指、打哈欠，并在交易大厅里闲逛。随后新消息出现了，可预测公司刚刚收到了一大笔意外的产品订单。听到这一消息，每个人都开始躁动不安，都试图尽可能多地分析这一消息。投资者更新了他们对可预测公司未来现金流和风险的预期，并提出新的估值，从而引发新一轮的交易。交易时刻都在进行，直到每个人都十分满意为止；然后，市场达到均衡并回归安静模式，直到下一条消息闪现。

正是市场消化信息的效率引起了价格以随机游走的形式浮动。如果

每个人都在吸收所有可用的信息并正确地进行计算，那么唯一会导致可预测公司股票价格浮动的因素就是新消息。由于传统经济学假设新消息的出现是随机事件，并且没人知道它是好消息还是坏消息，因此价格会随着消息的风吹草动而随机变动。但是，由于工业化世界的经济平均每年增长几个百分点，因此经济学家在模拟消息的随机性时，会偏向使用好消息。虽然不了解可预测公司的任何其他事项，但是公司在整体经济增长中是受益的，因此可以预计好消息比坏消息多一点，这被称为带漂移项的随机游走过程（random walk with drift）。虽然股票市场是随机游走的，但平均来看，随着时间的推移，股票价格仍会上涨。

拉斯维加斯、丘吉尔园马场和华尔街

随机游走假说的一个重要含义是过去股票价格变动中的信息无用。过去股票的涨跌情况不会影响未来的股票涨跌——再次强调，促使价格发生变动的唯一因素就是随机的消息。如果过去的价格变动有任何模式，理性投资者就会发现，这些模式会提供有用的信息，投资者在设定估值时就会运用这些信息。然后，他们利用模式从市场中套利，市场再次变得随机。

此外，理性投资者和套利的结合意味着打败市场并产生比整个市场更高的回报是不可能的，这被称为有效市场假说（efficient-markets hypothesis）。这一理论早期的构建者、芝加哥大学的尤金·法马指出，如果价格“完全反映”所有可用和相关的信息，那么可以称其为有效市场。[8]这意味着如果每个人都使用相同的公式来计算价格，并且可以获得相同的信息，那么价格将反映出这一点。既然无法比完全理性更聪明，那么只有拥有其他人没有的信息时，人们才能提出更好地反映股票“真实”价值的价格。现代市场中对发布和使用信息有着特别的管理，以防止公司职员或其朋友和家人等“内部人员”进行不公平的交易。当新的信息到达市场时，无论是公司的收益报告还是政府的经济统计数据，全天候的新闻报道以及互联网的普及实际上保证了每个人都能同时获得信息。有效市场的含义是没有办法获得优势并系统地击败市场。正如保罗·萨缪尔森曾经说过的，“在拉斯维加斯、丘吉尔园马场或是地方的美林证券公司，要想致富是很难的”。[9]

在这一点上，你可能会说，沃伦·巴菲特不就成功了吗？难道没有聪明的投资者和愚蠢的投资者之分吗？传统的金融理论家会回答说，是

的，有这种差异，但正是聪明的投资者设定了市场价格。市场中有两种投资者：一种是像沃伦·巴菲特一样精明的投资者，他们搜索所有最新的信息，并且运用智慧完美地分析信息；一种是平庸的投资者，他们根据出租车司机和调酒师的提示购买股票。当他们凑到一起进行交易时，精明的投资者会与平庸的投资者进行买卖，利用后者的无知，直到市场价格与精明投资者给出的价格相当并且各方停止交易为止。平庸的投资者会损失钱财，精明的投资者会大捞一笔，因此，随着时间的推移，精明的投资者掌控的资本金额将会增加，而平庸的投资者的资本将缩减，大部分资金留在了精明的投资者手中。[10]在真实世界的市场中，大部分资本都在养老基金、共同基金和其他专业投资者的控制之下，通常正是这些投资者进行的交易决定了价格。

平庸的人可能没有机会与职业选手对抗，但是随着时间的流逝，共同基金广告里的图表怎么样了呢？这些共同基金的广告总是展示那些回报高于市场平均值及竞争对手的特定基金。根据传统经济学家的说法，可能有两种解释。第一种解释是该基金承担的风险高于市场平均水平。在有效的市场中，人们可以获得比市场平均水平更高的回报，但是天下没有免费的午餐，人们付出的代价就是更高的风险。因此，在调整风险的基础上，人们仍然无法打败市场。第二种解释是运气。正如与“优秀”公司同行，如果在一定时期内拍摄一组共同基金的快照，就不可避免地会发现一些“优秀”的共同基金，就像一群人在屋子里掷硬币，总会出现许多硬币正面朝上的情况。正如马尔基尔在1973年出版的书《漫步华尔街》中所指出的那样，一些研究表明，随着时间的推移，在平均情况下，共同基金和专业投资者没有在具有统计意义并基于风险调整基础上超出如标准普尔500一类的广泛市场指数。事实上，他们中的大多数都在追随指数。[11]

传统金融将其理念发展成了严谨而有说服力的理论，这些理论可以一路将我们从个人投资者的行为一直带到全球市场的运作之中。传统金融理论在很多方面都是传统经济学中优雅和高度发达的分支之一。

然而，在过去的20年里，现实对这一理论并不友好。正如本章开头所指出的那样，诸如1987年的崩盘和20世纪90年代的技术泡沫等事件已经极大地挑战了传统金融理论。但也许更重要的是，使用先进的统计工具进行实证工作的新一代经济学家已经在质疑该理论的核心原则。从

这项工作的有利的角度来看，该理论充其量只是在某些情况下大致适宜；在最坏的情况下，我们可以说理论太简单而且是错的。为了切入问题的核心，我们需要明确三个显著的事实：

- 第一，大量的实证和实验证据表明，现实世界的投资者看起来并不像理论中描述的是完全理性的。[12]投资者不会像传统理论所假设的那样贴现，他们对风险存有各种偏见，在处理信息时会出现框架错误，并且会摸索着作出决策。正如第6章已经概述过的，这并不意味着投资者是非理性的或仅仅说是根据情感作出决定。相反，他们是“有限理性”和归纳理性的，而不是完全理性和演绎理性的。可以这么说，30年的工作毫无疑问地表明，完全理性没有体现出经济现实。
- 第二，巴舍利耶错了，市场并不是随机游走的。通过审视拥有较大结构的市场数据我们发现，它们具有复杂适应系统的所有典型特征。
- 第三，在传统经济学意义上，金融市场并不高效，但正如我们将要看到的，从进化的意义上来看，金融市场的效率非常高。

棉花价格、肥尾与分形

首先，为了理解传统理论出错的地方以及金融市场究竟是什么样的，我们需要回到20世纪60年代IBM公司位于纽约约克城高地（Yorktown Heights）的大型研究中心。当时，其中正崭露头角的明星之一便是在波兰出生、在法国接受教育的数学家贝努瓦·曼德尔布罗特。[13]他年轻的时候就是个叛逆的人。虽然法国数学界对布尔巴基学派（Bourbakist）的纯粹逻辑着迷，曼德尔布罗特却对几何学这一较冷门的学科感兴趣，通常会首先选择绘制数据图片而非书写公式。

1960年，曼德尔布罗特在哈佛大学做演讲，当时黑板上的一张图片吸引了他的目光。这张图片是经济学教授亨德里克·霍撒克（Hendrik Houthakker）收集的棉花价格数据图。亨德里克一直在绘制棉花价格的每日变化图，并试图将数据拟合成高斯钟形曲线。按照随机游走假设所说，亨德里克应该能达到目的，但是他无法获得适合的数据。然而，曼

德尔布罗特认为他发现了那些数据正在创建的模式，他把亨德里克的数据输入一个装有计算机（旧时数据编码中用的）穿孔卡的盒子里，带回了约克敦海茨。1963年，曼德尔布罗特在题为《特定投机价格的波动》的论文中公布了他的分析结果。他通过四个步骤瓦解了随机游走假说：（1）与钟形曲线相比，数据分布的厚尾特点非常明显。换句话说，价格波动比随机游走预测得更为极端。（2）极端事件非常极端，大部分的差异都受到少数猛烈的价格变动的影响；（3）价格变动似乎在时间上出现了一些聚集，换句话说，出现了一种间断均衡的模式。（4）与随机游走预测不同，描述数据的统计信息并不是静止的，而是随时间变化而变化的。[14]

曼德尔布罗特不仅否定了随机游走假说，还提出了另一种假说。他指出棉花价格分布的尾部遵循幂定律。[15]曼德尔布罗特的提议完整地解释了数据的肥尾和极端波动，所有这些特征都是随机游走假说无法解释的。在后来的工作中，曼德尔布罗特将金融市场价格描述为“分形几何”（fractal geometry）——不仅在金融数据中有这样的结构，这些结构还出现在从数分钟到数个月内的多个时间尺度上，就像因曼德尔布罗特而备受关注的分形图像的模式一样。[16]

1963年，曼德尔布罗特的论文发表，一石激起千层浪。随机游走假说的主要倡导者保罗·库特纳说：“要想推翻几个世纪以来的努力成果，我们首先得鉴别并确保曾经的努力是否真的一文不值。”[17]随后人们围绕曼德尔布罗特的论文展开了辩论，芝加哥大学的尤金·法马对此文进行了深度的批判。[18]但在1967年发表的另一篇论文中，曼德尔布罗特回应了批评者们，并用包括小麦价格、铁路公司股票、利率和汇率的额外数据证实了他的结论。[19]最终，辩论逐渐平息，经济学家们转向了其他话题。曼德尔布罗特是一个外行，是IBM公司的数学家，而不是经济学家。随机游走假说体现的是传统理论的力量，而在当时，没有其他理论支持曼德尔布罗特的结论。直到20年后，才出现了复杂适应系统以及理解该系统的工具。经济学家把曼德尔布罗特的论文视为“令人感到不适的异常现象”。曼德尔布罗特的论文同巴舍利耶早期的工作一样被遗忘了，等待着新一代研究人员重新发现。

华尔街的非随机漫步

1986年，麻省理工学院的一名叫罗闻全的年轻教授和他的合作者、

沃顿商学院的克雷格·麦金利，向一场庄严的金融学者会议提交了一份论文。两人在论文中宣称，他们已经证明股票价格并不是随机游走的。他们在1999年出版的《华尔街的非随机漫步》（*A Non-Random Walk Down Wall Street*）一书中描述了人们的震惊反应。[20]当他们继续探讨这一话题时，一位资深的经济学家言之凿凿地声称不可能出现这样的结果，还争辩着说他们一定犯了计算机编程方面的错误。

但在接下来的几年里，人们发现两人采用的程序合理，结论有理有据。与曼德尔布罗特不同，罗闻全和麦金利是金融领域的行家，他们在研究中使用的是传统金融研究的主流技术。在两人发表论文之前，随机游走假说的统计测试已经进行了几十年，但测试的重点是是否可以根据过往的股票价格预测未来股票的价格。[21]测试结果是否定的，但正如罗闻全和麦金利指出的那样，虽然测试结果与随机游走假说一致，却不能证明这一假说是正确的。他们提出了一个略有不同的问题：数据中存在的任何模式或规律是否会以统计意义的方式偏离纯粹的随机模式？[22]而人们是否可以使用任何此类模式来赚钱是另外一回事。两人发现了两种能够否定随机游走的模式，这两种模式都符合曼德尔布罗特的发现。首先，他们发现价格和时间之间存在相关性；其次，市场价格的变化与随机游走假说所描述的不同。[23]

在接下来的10年中，罗闻全和麦金利大部分时间都在回应各种针对他们的研究结果的批判，并进一步寻找证据来支持研究结果。此外，其他研究人员也配备了越来越强大的计算机和统计工具，开始探索世界各地金融市场数据中各种有趣的模式，[24]随机游走假说迅速消亡。到2000年《华尔街的非随机漫步》第七版发布时，作者伯顿·马尔基尔也不得不承认实际上市场并不是随机游走的。

来自经济物理学家的攻击

正当罗闻全和麦金利从行内人的角度攻击传统金融理论的基础时，另一组研究人员开始从外部进行攻击。1989年11月9日，柏林墙倒塌，突然间，所有人都在谈论国防开支的重大削减和“和平红利”。对于在洛斯阿拉莫斯（Los Alamos）和桑迪亚（Sandia）等美国国家武器实验室工作的数千名物理学家和数学家来说，“和平红利”却意味着失业。在世界的另一端，苏联的世界级科学家也面临失业。同时，蓬勃发展的华尔街也急需训练有素的数学家。很快，需求和供给得到了匹配，世界各地

的大量火箭科学家开始在高盛、摩根士丹利和各种秘密的对冲基金就职。科学家之间的竞争依然激烈，只不过竞争的重点从炸毁世界转为价格衍生品。

起初，物理学家只是简单地消化了别人给他们的财务教科书，并开始研究计算机模型找方法赚钱。许多人都注意到了他们所看到的现实数据与教科书之间的差异。与此同时，混沌理论、非线性动力学和复杂性理论在物理学领域取得了突破性的进展。终于，物理学家转向这些理论寻求答案，而金融和复杂性理论最终相遇。很快，传统物理科学期刊，如《物理评论快报》（*Physical Review Letters*）和《自然》杂志中印满了由“经济物理学家”撰写的金融文章。[25]到20世纪90年代，基于曼德尔布罗特的主张，尤金·斯坦利（Eugene Stanley）、琼·菲利普·布沙尔（Jean Phillippe Bouchard）、尼尔·约翰逊（Neil Johnson）、罗萨里奥·曼特尼亚（Rosario Mantegna）和迪迪埃·索内特（Didier Sornette）再次活跃起来，并开始用波动、肥尾的股票市场这一视角对抗传统金融学，这与教科书的视角截然不同。[26]

多因·法默正是一位从武器实验室转型进入华尔街的物理学家。法默是原圣塔菲研究所经济学研讨会的成员，本书第8章回顾了他关于幂定律和股票市场的一些研究。1991年，法默辞去了洛斯阿拉莫斯国家实验室复杂系统小组负责人的工作，与他的老朋友物理学家诺曼·帕卡德（Norman Packard）创建了预言公司（Prediction Company）。新公司的使命是将最新的物理理念应用于金融市场。作为物理界的知名人物，法默不仅因其在复杂系统中的开创性工作而享有盛誉，年轻时他还尝试将物理知识用于轮盘赌来决胜拉斯维加斯。他和包括帕卡德在内的几个同伴，将轮盘赌的非线性模型写入了小型计算机，又把小型计算机装进鞋子里，然后出发前往赌场。他们并没有赚到预期的数百万美元，但最终大获全胜，因此引发内华达州颁布禁止计算机辅助赌博的法律。预言公司就像“鞋子里的电脑”，是法默和帕卡德用来在世界上最大的赌场——华尔街放手一搏的工具。法默在其后的8年中在真实世界的市场工作，并与包括瑞士联合银行（United Bank of Switzerland）在内的客户进行交易。后来，法默决定回到学术界和圣塔菲研究所，将其所学应用于理论世界。[27]

市场是不断进化的生态系统

传统金融学的一个基本主张是市场中的任何模式或信号都会被警惕和贪婪的投资者用来套利。如前所述，传统金融学假设所有投资者都可以获得相同的信息，如果股票价格存在任何模式，投资者将会发现这些模式，并在定价决策中将其考虑在内，从而推动市场回归随机游走。然而，法默发现传统经济学中对于套利的定义受到第3章提到的20美元钞票问题的困扰。传统金融学中的均衡框架导致经济学家忘记了时间，也忘记了市场是动态系统。法默和他在预言公司的同事发现，他们可以在市场数据中发现有统计意义的信号（见图17-1）。[28]这些“信号”包括复杂的模式和预测未来股价的各种因素之间的关系，例如利率、交易量——由于这些信号是私有的，他们并没有详细披露其组成部分。传统理论认为，任何此类信号一经发现，会被立刻用于套利。但法默和他的团队发现，这些信号会持续一段时间，通常为数天或数月，有时甚至持续长达10年（见图17-1a）。随着交易者发现并利用这些信号，他们常常会看到信号随着时间的推移而减弱。但他们还发现，市场的复杂非线性动态意味着即使旧信号由于套利而消失，新信号也在不断被创造（见图17-1b）。

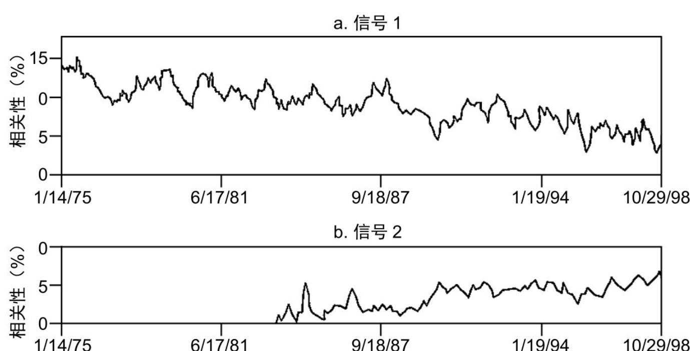


图17-1　信号出现并消失：1975—1998年两种私有预测信号的强度

资料来源：法默（2001年）。

法默从交易经验和研究中获得的重要信息是，市场形成了一种不断进化的生态系统。市场上有各种各样的交易者和投资者，他们都有不同的心智模式和策略。随着时间的推移，这些主体会相互影响，不断地学习和调整策略。可以说人们一路在所有可能的投资策略库中跌跌撞撞地寻找方向。这些主体复杂的互动，不断变化的策略，随着时间的推移出现的新信息都会催生或者消灭模式及交易机会。

圣塔菲研究所的布莱恩·阿瑟曾诗意地称市场为“预期的生态系统”来

描述主体与其策略之间的相互影响。[29]在第6章中，我们讨论了由阿瑟、约翰·霍兰德及其同事建立的圣塔菲人工股票市场模型。[30]他们用于替代传统金融学完全理性的方案便是投资者具有“归纳理性”。[31]在圣塔菲人工股票市场模型中，主体参与了一个进化搜索，寻找有利可图的策略。关于哪种策略会在特定时间取得成功取决于市场中其他主体使用的策略。随着一种动态被创造出来，其中特定的策略组合将在市场中创建模式或结构，它们会在其他主体寻求利用这些策略时改变其他主体的行为，这会进一步促生其他模式，导致其他主体作出反应，等等。圣塔菲人工股票市场模型的结果很好地复制了现实世界市场的关键统计特征，例如集群波动（clustered volatility，即间断均衡）。此外，该模型定性地复制了法默观察到的模式来和去的情形。[32]然而，该模型有一个局限性，即它的现实主义倾向使它变得更复杂了。事实上，主体可以从近乎无限多的可能策略中进行选择，这意味着该模型本身几乎与真实的市场一样难以理解。基于物理学方面的研究，法默也知道复杂的宏观行为有时可能是非常简单的微观行为的结果。正如法默所说，他的目标“不是建立最接近实际情况的可能市场模型，而是建立最简单且合理的模型”。[33]

根据真实市场中的交易经验，法默开始研究一种新的、更简单的模型。[34]他并没有让每个主体拥有“个性化”行为，而是假设市场上只有三种类型的投资者。[35]第一类是价值投资者，他们通过查看公司的收益、增长、竞争地位等信息信号，基于所谓的基本面买入或卖出股票。实际上，价值投资者正是按照传统理论的设定行事的。

第二类是技术交易者。这些投资者不考虑基本面，而是考虑过去的价格和交易量来确定策略。例如，他们可能会看到一支股票的价格开始上涨，然后买入，赶上涨潮。传统理论认为，这些投资者永远都不可能获利，因为如果市场随机游走，过去的价格其实没有任何意义。然而，正如法默指出的那样，现实世界充满了运用技术或混合技术价值策略进行交易的人，因此，重要的是看这组投资者可能对市场产生什么影响。

法默随后增加了第三种主体，即流动交易者。例如，这种人可能会购买房子并出售一堆股票以获得首付金。这样的交易者不卖出股票，因为他们认为股票价格被高估或低估了，他们只需要流动资金。

随后，法默在他的人工股票市场中放入了一名终极主体。该主体不是交易者，却发挥了关键作用：做市商。大多数传统的金融模型都预先假定现实世界股票市场如何运作的制度细节并不重要，价格是由瓦尔拉斯拍卖会设定的。[36]正如前文所述，瓦尔拉斯引入了标准的传统假设，

即像神一样的拍卖师能够匹配买家和卖家以设定价格。瓦尔拉斯关于拍卖商的灵感来自19世纪末的巴黎证券交易所（Paris Bourse），而大多数现代金融市场很久以前就从瓦尔拉斯的拍卖会过渡过来了，它们现在采用的是连续的双向拍卖方式（continuous double auctions，连续是因为拍卖在市场开放的整个时间段运行，双向拍卖是因为交易者可以买入和卖出）。正如我们之前看到的，法默和他的同事们已经证明，市场的制度细节确实在市场动态中发挥了重要作用，尤其是产生波动的幂定律分布。与真实世界的市场相比，虽然这个特定模型中的做市商仍然是简化后的版本，但仍然比典型的传统模型更接近现实，因为它使用的是连续的双向拍卖方式，而不是瓦尔拉斯式的拍卖。这点很重要，因为这样可以确保将交易的时间动态和大订单的价格影响考虑在内，而瓦尔拉斯拍卖方式则假设一切都是同时发生的。

价格不等于价值

最初，法默创造了一个只有一个基本主体和做市商的市场。随后他进一步简化市场，方式是假设基本的主体始终能够确切地了解股票真实、完全的理性价值。为向传统金融学致敬，法默也假设真正的价值是随机游走的。然后他给基本主体设定了一个简单的交易规则：如果股票的价格低于基本价值，买入；如果股票的价格高于基本价值，卖出。而法默的期望是，这个简单的设置拥有一个完全理性的主体和完美的信息，会复制传统金融学中达到均衡的条件，同样，市场价格将追踪真实价值并随机游走。

然而事实并非如此（见图17-2），股票价格与其价值大致相当，但也并不完全匹配。主要的原因是存在时间延迟，同我们在第5章动态系统的讨论中检验的时间延迟性质一样。这里可以回想一下淋浴器的例子，旋转旋钮和水温变化之间存在延迟——首先你会在理想温度附近摆动，然后找到合适的范围，最后摆动幅度变得越来越小，直到达到目标温度。我们假设并不存在恒定的理想温度，理想温度会随机波动。在这种情况下，摆动永远不会完全消失。此外，当理想温度大幅上升或大幅下降时，你和理想温度的差距变得更大，就需要不断地追赶它。最终，你会赶上它，甚至稍微超过它。因此，随着时间的流逝，淋浴器的实际温度将大致追随理想温度的随机游走路线，但并不完全匹配。你在淋浴时的体验看起来就像图17-2中展示的情况。在法默的人工股票市场模型中，交易者的行为和做市商的反应之间存在着短时间的延迟，因此两者

之间产生了动态关系，导致股票价格和价值之间并不完全匹配。传统金融的均衡观点只是假设一切都是同时发生的，因此隐藏了这些动态。

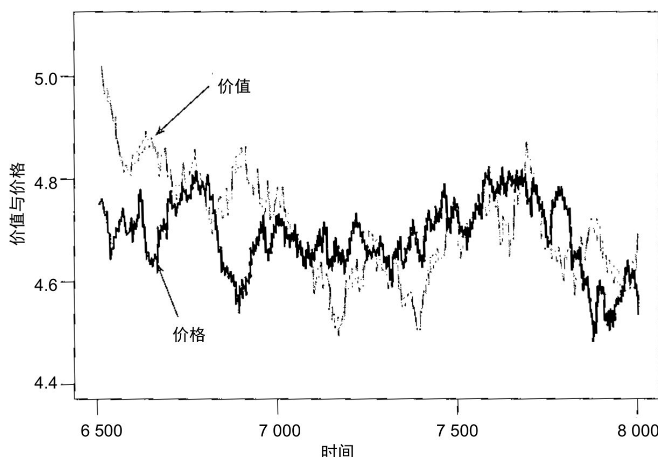


图17-2　理性主体前提下的价格与价值对比图

资料来源：法墨与乔西（2000年）。

一位传统经济学家可能会说："很好，但股票价格与价值的偏差是不偏不倚的，所以可以视其为随机噪音，而实际做市中所涉及的时间延迟非常短暂，只有几秒钟或者几分钟，所以从几天和几周的时间维度来看，这真的不重要。"但是也有人提出了不同意见。首先，偏差本身并不是随机的。虽然真实价值可能是一个随机数，但价格取决于交易者与做市商的互动——它不是噪音，而是物理学家所说的时间结构。也就是说，交易者和做市商之间的互动形成的价格本身具有趋势和动力。其次，法默知道，虽然这个时间结构可能会在秒和分钟的层面上起作用，但这一结构的作用可能像牡蛎中的一粒沙子一样，随着时间的推移，通过系统的动态形成更大规模的模式。

为了证明这一点，法默将技术交易者放入了模型中。为了简化起见，他再次假设技术交易者是一个趋势追随者：当股票价格上涨时，交易者买入股票；当股票价格下跌时，交易者卖出股票。将技术交易者加入模型之中的效果是动态增强。回到淋浴的例子，假设再次淋浴，你会尝试将实际温度与随机波动的所需温度相匹配。但是在地下室，还有其他人转动供水的旋钮（技术交易者）。现在，你打开热水时会发现有一点延迟，但最终温度会开始上升。同时，地下室的技术交易者看到温度上升就会转动控制旋钮，在系统中加入更多的热水，助你一臂之力。接着，你开始加入凉水来维持水温，就像在与技术交易者较劲。你加入了

足够的冷水，温度达到峰值后开始下降——但是地下室的技术交易者还是在和你对着干，他开始向系统中加入冷水。这样你就很难将实际水温与理想水温相匹配。最初的微小波动现在已经变成了大幅波动，在很长的时间内，水不是太热就是太冷，甚至会变得滚烫或结冰。

这实际上正是法默模型中发生的事情。加入追随趋势的技术交易者正是关注到价格中的时间结构并加强了这一状态，在基本面和技术交易者的互动下，出现了明显且长期与真实价值偏离的情形。[37]法默发现其他真实的世界因素进一步加强了价格和价值的分离。特别是，成功的投资者倾向于对收益进行重新投资，从而增加资本并提升其战略对市场的影响。如果技术交易者在一段时间内获得成功，其资本增加将会进一步推动价格上涨，偏离价值。

在这一点上，传统经济学家可能会控诉："套利者难道不会将价格推回基本价值吗？"根据传统金融学理论，价格和价值可能会不同，但两者的差异只是随机的短时噪声，因为价格会在真实价值附近波动，因此，平均而言，价格等于价值。由于偏离基本价值是随机的、是在短期内出现的，因此没有人可以从中赚钱，唯一合理的策略是成为一个基本面投资者。这一假设强调了有效市场假说中的一个悖论。米尔顿·弗里德曼首先指出了这一悖论，即市场需要技术交易者从偏离的基本价值中套利并保持市场效率，但技术交易者无法在高效的市场中赚钱，因此会离开市场。这一情况也引发了一个几十年来一直困扰经济学家的问题：如果市场如此高效，而且技术交易者无利可图，那么为什么会存在那么多技术交易者呢？

法默找到了这一悖论的答案。他采用自己的模型，并将股票的基本价值设定在一个恒定的水平。然后，他创建了一组主体——季节性交易者，他们只是在可选择的模型中买入并售出股票。他们的行为在股票价格中创造了一种简单、规范、波动的模式。接下来，法默引入了一批技术交易者。根据过去一段时间内股票价格上涨或下跌的情况，他给每位技术交易者设定了一个随机生成的策略。每个主体都会获得一笔启动资金，成功的交易主体会保留利润并再次投资。这样就形成了一种简单的进化系统。主体会尝试各种技术策略，然后重复有效的策略，摒弃无效的策略。如果传统金融学是正确的，最后应该会有一方大获全胜。一旦采取正确策略的技术交易者自动锁定波动模式，他们进行的交易将抵消波动，从季节性交易模式中套利并将其排除在市场之外，市场再次恢复高效。

法默在运行他的模型时，发现了一些更有趣的事情。起初，一切都与传统理论相符。技术交易者刚开始时没有多少钱，所以他们对价格并没有太大的影响。但很快他们就注意到了波动模式，开始套利，赚了很多钱。技术交易者越来越成功，开始进行更大规模的交易，这反过来又开始影响价格。经过一段时间后，随着交易者从低效率的市场模式中套利又将其排除在市场之外，法默发现波动模式减弱，进而价格更接近基本面价值（见图17-3）。经过5000个周期后，波动几乎消失了，市场看起来非常接近完美。但随后，波动性突然爆发，价格开始变得混乱。这一过程中发生了如下事件：技术交易者越来越富有，他们的交易量变大，大型交易开始将本身的动向引入价格中。这些动向为其他技术交易者创造了机会，他们试图从其他技术交易者创造的模式中套利——当技术交易者在季节性交易中获利后，他们开始互相喂食！

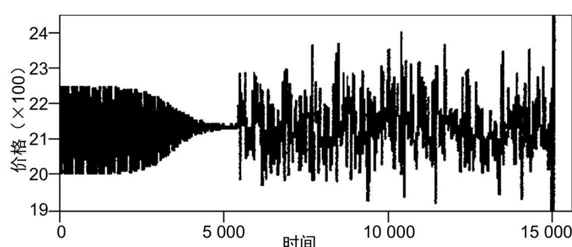


图17-3　基本面交易者无法从技术模型中套利

资料来源：法默（2001年）。

法默继续运行他的模型，进行成千上万次的迭代，这种疯狂的模式不断地向前推进。图17-3中显示，即使运行5　000个周期之后，整个模式看起来仍是随机的，但事实并非如此。该模型是完全确定的，根本没有随机性，紧急模式完全由交互代理的动态产生。一位传统经济学家可能会反驳称，技术交易者之间的行动是一场零和博弈，钱只是从一位技术交易者的口袋转入另一位技术交易者的口袋，平均而言，没有人真正从中赚钱。如果人们是完全理性的，就不会玩游戏。他们会在5　000个周期之时退出，市场会恢复原状，达到有效的平衡。

但正如之前所讨论的那样，有证据表明，在实际市场中，人们并非完全理性。相反，人们偏向于保持乐观，相信他们可以比下一个人做得更好，也会随着事态发展不断学习。随着投资者在市场中尝试不同的策略，他们会在清醒之时离开，而其他投资者会注意到并利用这种模式。当这些其他投资者利用这些模式时，结果就是出现了新的模式，其他人仍然会接受这种模式，然后依次向前推进。一些投资者暂时赚到了钱，

但随后他们又会赔钱。新的投资者看到机会时会进入市场并引入更多资本，一些投资者则会在策略失败时退出。法默的模型展示了他在现实交易经历中目睹的种种行为。随着市场的复杂动态展开，时间一长，模式循环往复，市场在均衡和风暴间摇摆。

法默的模型不仅与他个人的经验相符，还与大量的实验证据相一致。首先，法默的结果复制了现实世界数据的关键定性统计特征，包括曼德尔布罗特发现的肥尾特征。其次，许多研究表明某些形式的技术交易者可以盈利。[38]最后，一些研究表明，基本价值只能解释整体股票价格变动的一小部分。图17-4b中展示了一项此类研究的结果，很明显，此结果与法默模型（见图17-4a）的研究结果非常相似[39]：价值和价格之间的巨大差异可以持续多年，甚至几十年。

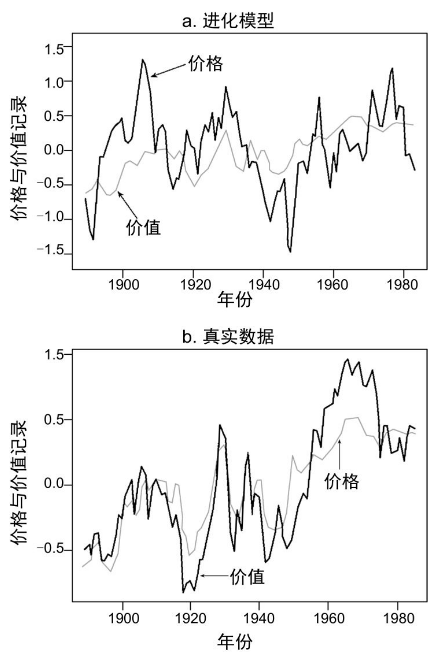


图17-4　进化模型在质量上与实际数据相似

资料来源：法默（2002年）。

市场效率的新定义

金融市场是否有效率？从复杂经济学的角度来看，金融市场中充满了相互竞争的交易策略，是一种不断进化发展的生态系统。因此，询问市场是否有效与询问亚马孙雨林的生态系统是否有效一样重要。那么，

衡量效率的参照物是什么呢？

我们还可以提出一系列更有意义、更具体的问题：市场对新信息的反应速度如何？答案是非常快。有大量研究表明，当信息出现时，市场几乎会立即作出反应。[40]市场进化系统的竞争力确保了这种敏锐性，正如亚马孙雨林中一株树的树冠上掉下一片柔软的嫩叶，一大堆昆虫几乎立即就会冲上去。事实上，随着美国有线电视新闻网络、彭博社、互联网和大型计算机性能等的发展，市场处理信息的能力变得越来越快。传统金融学理论预测，获取更多的信息后，价格相对于价值的波动应该减少，而实际情况刚好相反。随着信息和技术力量的拓展，似乎市场变得更加不稳定了。[41]进化理论对此的解释是，信息技术的进步拓展了构思投资策略的空间，增加了套利和技术交易的机会，因此我们会看到所谓的对冲基金数量急剧增加。[42]

市场是否能够有效地处理信息？答案是肯定的。市场能够对从人群中获得的大量信息进行分类，并且综合考察这些信息的含义。[43]已有实验利用市场来预测选举结果以及体育比赛的比分，以及奥斯卡颁奖典礼的结果等方方面面的内容，并且市场经常作出更多的预测，这些预测结果比个别专家的预测或者民意调查更精准。[44]这种利用信息的效率是有效市场假说的基石。然而，该理论假设市场的强大预测能力只用在与股票基本面相关的信息，而在现实世界中，交易者也会猜测其他交易者正在进行及思考的事情，并对此下注。在这种动态更强的情形下，市场中的人们在不停地产生预期，又在不断地揣度别人的预期，市场处在一种无限循环的预期的环路里。约翰·梅纳德·凯恩斯就曾指出了这种无限回归——“人们对舆论的走向有所预期，我们用智慧来预测这种预期”。布莱恩·阿瑟在他的酒吧问题中详细地从数学角度阐释了这一问题。[45]那么，这些对期望的预期是否满足了市场的进化动力？答案是没有，这意味着虽然市场是非常强大的信息处理方，但在传统经济学看来，市场的效率并不高。

投资者是完全理性的吗？不，但投资者是利己主义的。金融市场不是慈善机构，而是受到恐惧和贪婪的驱使。投资者聪明吗？总的来说是这样的，大多数金钱都是由专业人士管理的，包括那些聪明的街头交易者以及火箭科学家和博士。资金经理人拥有人类的所有优势，例如精湛的模式识别能力和创造性的学习能力，他们也拥有人类所有的脆弱性，包括非理性和偏差。市场竞争激烈吗？答案是“是的”。看看交易大厅里那些焦虑不安的面孔，你就能清楚地看到这一点。随着全球资本流动的

管制放松，市场的竞争越来越激烈，也只会越来越激烈。

聪明、有竞争力的投资者能否从有利的机会中套利并推动市场提高效率？有时候是这样的，但并非总是如此。法默指出，传统理论对市场效率作出了两个关键的假设。[46]第一，为了实现套利，交易者必须能够识别出产生异常利润的策略。第二，他们用于实现异常盈利策略的资金数量必须足以套利并使市场恢复效率。法默指出，这两种假设放在现实世界中都存在问题。衡量一个异常盈利策略的标准是信息比率，也称夏普比率（Sharpe ratio）或回报风险比率，即预期年回报率除以该回报的预期标准差。高回报、低风险的策略具有较高的信息比率，而低回报、高风险的策略具有较低的信息比率。大多数交易者认为，信息比率为1时的策略比较诱人。但正如法默所说，了解策略的信息比率需要花费时间。假设收益呈正态分布，那么根据统计法则，需要4年的数据才能确保交易者采取的策略信息比率为1（但由于数据挖掘本身的问题，我们又不能仅依靠历史数据，必须实际测试未来的策略）。如果收益具有肥尾特点，则需要更长的时间。因此，交易者通常采取的经验法则是，用5年的时间来了解一项策略是否有利，这可能是正确的，而这个时间尺度被“第一百万只猴子”的问题进一步延长了。如果有足够多的交易者进行交易，那么很可能他们当中有一个人运气很好，在5年的时间内都运转良好，这样识别异常盈利策略的时间又会延长。

第二个假设也存在类似的困境。传统理论假设有无限量的资本可以立即用于验证异常盈利策略。但在现实世界中，成功的交易者需要随着时间来积累资本。如果交易者一开始拥有100万美元并获得25%的回报，他就需要用30年时间将管理资金增加到10亿美元。如果这位交易者每年额外借助外部力量将其资本翻倍，则仍需要10年时间将资金增加到10亿美元。成功的交易者有可能更快地从外部投资者那里筹集大笔资金，但这仍然需要几年时间来建立业绩记录并筹集资金。但是在这几年的周期中，市场风云变幻，因此盈利的交易策略从出现到消失的时间跨度可能会比从市场中套利又恢复效率的时间要长。

是否存在神奇的致富公式？并没有。萨缪尔森说过，在拉斯维加斯、丘吉尔园马场或是当地的美林证券公司，要想致富都是很难的。他说得对，但是他给出的原因是错误的。并不是因为金融市场跟随随机游走，相反，无论是华尔街交易者，还是巴西雨林中的树蛙，都很难在任何竞争激烈的环境中变得富有。不过，华尔街和拉斯维加斯之间有一点不同。在拉斯维加斯，无论你怎么努力，无论你有多聪明，都无法提升

致富的概率，除非在鞋里安装一个计算机。但在华尔街，致富并不是只靠运气——这里是高度活跃、不断进化的复杂系统。在任何时候，总有些人对市场的了解程度比其他人高，并且至少会在一段时间内发财。传统经济学需要反思的是，如果金融市场如此高效，那么为什么华尔街上停着的保时捷和法拉利比其他地方更多呢？但金融市场中的竞争优势稍纵即逝，产品和服务市场也是如此。总有人有更好的想法，采用更聪明的策略，或者拥有更强大的技术，从而脱颖而出。

有效市场假说是19世纪的均衡理论和巴舍利耶随机游走理论之下诞生的神话。虽然传统市场效率可能不是非常有意义的概念，但金融市场是高效的进化系统。市场是最好的社交技术，用于整合大量观点，为复杂资产定价，并分配资本。此外，多个市场的竞争强度可以确保市场能够快速处理信息，并使参与者处于压力之下，不断创新。麻省理工学院的罗闻全将市场的这种进化效率称为“适应性市场假说”。[47]如果说纽约华尔街、伦敦金融城或世界范围内的其他市场都有一个神奇的公式的话，那就是变异、选择和放大，并周而复始重复这一过程。

复杂经济学中有关金融的观点仍在发展之中，许多问题仍然没有答案。但很显然，传统理论存在缺陷，我们需要新的方法。华尔街、伦敦城金融市场和其他金融市场才刚刚感受到金融复杂性的影响。许多公司，包括银行和对冲基金，都在使用复杂经济学的统计和建模技术设计投资策略——许多公司雇用大量经济物理学家和行为经济学家或者听取他们的建议。当然，这些公司对自己正在做的事情以及取得的成果是保密的。然而，毫无疑问，复杂性科学将改变投资世界，并将在投资战略的“红桃皇后竞赛”中进行创新。

复杂性科学对个人投资者的影响可能是长期的。复杂性金融学中，个人投资者仍然能够得到宽泛的指导，例如多元化投资组合或长期投资非常重要。但是，大多数金融创新最终都会从金融机构投资者转向小额投资者，而且在某个时间点上，个人投资者也能够获取运用这些技术的资金。

在另一个领域内，复杂性金融学为企业高管提出了几个重要问题。复杂性融资在三个方面影响着商业世界。首先，广泛使用的计算公司资本成本的方法可能是错误的。其次，复杂性金融学质疑高管薪酬中包含股票期权的做法是否恰当。最后，也是最重要的一点，复杂性金融学给

人类带来了挑战，我们必须用新方式思考企业的根本目的和股东资本主义的本质。我们将依次讨论这三个问题。

资本成本

公司的资本成本听起来相当平常，更像是技术性话题。但在企业高管作出一些最重要的决策的时候，包括要进行哪些投资，要采取哪些策略，以及是否进行合并或收购时，公司资本起着至关重要的作用。[48]

所有公司都要从某处获得资金，无论是银行贷款还是股东投资。该资本的提供者需要获得回报，而资本成本只不过是各种资本提供者预期收益的加权平均值。如果公司一半的资金都来自银行贷款，利息为6%，而另一半来自投资人，他们预期的收益率为12%，那么该公司的资本成本将达到9%。这一数字为公司提供了一个基准——管理层作出的任何投资决策必须提供至少9%的回报，以支付所用资本的成本。由于股东承担着管理层是否成功完成这项任务的风险（债务持有人因受到保护而不会受到损害，但他们也无法分享上行空间的利润），因管理层决策产生的任何超过资本成本基准的回报都会为股东创造价值；而当回报低于资本成本基准时，就会损害股东利益。

判断管理团队运营是否良好时，计算出基准数值就显得非常重要。然而不幸的是，计算这个数字时使用的最广泛的方法直接来自传统金融学理论。标准的方法被称为资本资产定价模型（capital asset pricing model），这种标准方法中包含所有常见的与完全理性投资者、有效市场和均衡相关的假设。[49]这种方法还对投资者管理风险的方式作出了重要假设。20世纪50年代，芝加哥大学经济学家哈里·马科维茨提出了资本资产定价模型，这一模型假设所有投资者都持有股票组合，能够充分降低风险、提升收益。如果市场中的每个人都拥有这样的投资组合，那么他们就可以相互结合创建市场投资组合，然后根据理论市场投资组合的相关性来测量单个股票的风险。这种风险因素随后被用于计算权益资本成本，或是计算为补偿股东承担的风险需要提供的回报。

问题在于，市场中没有人拥有马科维茨提出的完美投资组合，更不可能出现所有人都有完美的投资组合。[50]马科维茨的想法几乎不可能实施。例如，该理论假设人们完全了解公司的风险信息，可以无限制卖空股票，并且所有投资人都处在同一投资周期。此外，由于风险和收益属性会发生变化，市场组合必须不断更新。但在实际情况中，添加、删除

和重新加权投资组合中的股票涉及重大的交易成本，这意味着人们不会照理论所说的那样多次重新分配其投资组合。事实上，即使是经济学家本人也无法创造理论市场中的投资组合，当他们检验理论时，只是使用市场中所有股票的加权平均值作为代理。[51]在实际情况中，人们会使用更容易计算的指数，如标准普尔500指数或伦敦金融时报100指数来评估投资经理，而不是根据马科维茨理论中的市场投资组合来进行评估。而且，有证据表明，管理者“追逐回报”不是以资本资产定价模型中假设的方式优化风险与权衡收益的。[52]

资本资产定价模型中不切实际的假设意味着理论预测在实证检验中表现不佳，这不足为奇。最重要的是，我们已经可以证明作为度量风险的变量在衡量风险时几乎没有价值。[53]因此，基于资本资产定价模型公式计算公司的权益成本并没有什么意义，在现实世界中使用资本资产定价模型需要进一步的假设和判断，这更是雪上加霜。例如，计算股权较之于债务的风险溢价的变化很大，取决于所使用的特定指数和周期，并且必须对运用的无风险利率作出假设。[54]而从根本上来说，资本资产定价模型作为计算资本成本的方法，在理论和实践方面的价值都令人质疑。也有人提出了其他替代方案，例如“多因素模型”（有时被称为套利定价模型），但事实证明，这些方案与资本资产定价模型一样，有着许多共性的问题。[55]

沃伦·巴菲特总结了他对资本成本的看法：“我和查理·芒格一点也不清楚我们的资本成本是什么，我们认为整个概念都相当疯狂。坦率地说，我见过的资本成本对我来说都没什么意义。”[56]虽然传统的计算资本成本的方法存在缺陷，但是更好的方法还有待开发。[57]复杂性金融学研究计划中的优先事项就是开发一种替代资产资本定价模型的方法，以及一种衡量风险的新方法。目前，许多首席财务官和其他高管都坚持使用传统方法，但他们应该有保留地使用这些数字，并且做任何决策时都要彻底测试决策敏锐度对这些计算结果的影响。

股票期权是否有意义

20世纪70年代的学术金融革命的一个结果就是，董事会和管理团队衡量公司成败的主要指标成了股票市场表现。这反过来又掀起了一种通过使用股票期权将高管薪酬与公司在股票市场的表现联系起来的运动。

这场运动背后的动机和出发点非常有意义。公营公司受到内部的委

托代理问题困扰，这使得公司的股东常常与首席执行官及其管理团队相对抗。一方面，股东希望公司管理层能够将股份价值最大化。另一方面，首席执行官和管理团队可能会在管理过程中将其他部分最大化，比如他们的工资、大厅里的花哨艺术品和公司的规模。20世纪60年代和70年代，许多公司都受制于其管理团队，在首席执行官的主导下，公司状况极其糟糕。公司管理团队会最大限度地提高他们的薪水，自我膨胀。传统金融学教授提醒我们，股东才是公司的所有人，首席执行官和管理团队是为股东工作而不是为自己工作，这样每个人才能各司其职。

一个符合所有者和管理者利益的合理方式是衡量管理团队对股价所做的贡献，然后给予他们经济激励，以期提升股价。有观点认为，实现这一目标的最佳方式是为经理人提供股票期权，将他们未来的薪酬与未来的股票表现挂钩。这种想法在20世纪80年代风靡起来，像一场野火一样扩张，在一波恶意收购让董事会面临压力以提高其公司的股价时尤其如此。1983年，美国首席执行官薪酬中位数中，股票期权占至23%，而1998年上升到了45%。[58]

从理论上看，这是个好主意，但从复杂性金融学和实际实践的角度来看则有一点问题。大多数国家法律都规定了管理团队要为股东利益服务，具有信托责任，大多数复杂经济学家都会认同委托代理问题的重要性。同时他们也会指出一个重大问题：首席执行官和管理团队实际上并不控制股票价格。他们会控制战略、成本、投资和人员等因素，而这些会影响公司的收入、利润、资本回报和增长等。首席执行官和管理团队掌控的杠杆能够影响公司创造经济价值的能力，但只会间接影响股价。

传统经济学家认为，将管理人员的表现同股票价格捆绑在一起的做法还不错，因为在有效的市场中，从长远来看，管理层影响的基本面也会推动股价上涨。但正如我们所看到的，在实际当中，价格并不总是与价值正相关，而且偏差可能很大并持续多年。当然，价格和基本价值之间存在一定的相关性，市场对管理团队的行为并非完全脱敏，但这种相关性也并非密不可分。实际上，首席执行官正握着松动的方向盘向前行驶，而我们则是在这种情况下判断他驾驶能力的优劣。

有时候，投资策略之间的相互作用、交易者的心理、市场制度结构和其他超出管理团队掌控范围的因素对股价的影响远远超过基本面因素。例如，20世纪90年代股市出现巨大的泡沫，许多首席执行官一夜暴富，他们不过是顺势而为。同样，泡沫破灭时也产生了相反的效果，首席执行官的薪酬介于明星和普通职员之间。麦肯锡公司的研究表明，从

1991年到2000年，个体公司70%的收益与市场因素有关，只有约30%与公司特定的因素有关。[59]从理论上讲，人们可以消除大范围市场指数中的变动影响，但是它们与造成价格变动的主要因素的相关性是多少呢？我们目前仍无从得知。股票价格上涨是由于市场青睐首席执行官的新策略，还是因为几家大型对冲基金的计算机被锁定在一场多维数学战斗之中？结论不得而知。

将首席执行官的薪酬与股市表现捆绑也不可避免地改变了首席执行官的所作所为。在此影响下，有些首席执行官在责任感的驱动下，会更注重公司在股市中的表现。但是也有许多首席执行官执迷于市场的风吹草动，变得草木皆兵，在管理中目光短浅，仅考虑市场的需求（通常是下一季度的收益），而不会放长线、钓大鱼，不会专注创造长期价值，也不会建立强大的适应性组织。我们还看到了股市执迷于安然公司、世界通信公司和泰科公司（Tyco）的丑闻中，而管理层的恐惧和贪婪造成了收益操纵和彻头彻尾的欺诈行为。

这并不意味着首席执行官和管理层对企业的股市表现没有影响，他们当然会发挥作用，但也只是决定股价的众多参与者中的一员，而他们产生的基本面信号往往会被复杂多变的市场噪音淹没。我也不否认首席执行官与企业股票之间的联系，我们当然不想回到过去那种不负责任的首席执行官持有股东抵押物的黑暗时光。但事实上，我们不应让首席执行官对无法控制的股价负责，而应该让他们对实际掌控的事情负责，即获取基本的经济效益并创造长期价值。自20世纪90年代初以来，哈佛大学商学院教授罗伯特·卡普兰（Robert Kaplan）与戴维·诺顿（David Norton）共同写过一本书，也一直在倡导一种衡量公司业绩的方法——平衡计分卡（balanced scorecard）。[60]卡普兰和诺顿指出，没有任何单一指标，包括股价，能够完整地展示公司的业绩。相反，卡普兰和诺顿建议每个公司对业务模型中推动创造价值的因素进行分析，并根据自身特色围绕这些因素构建计分卡。计分卡应该包括会计、金融市场和其他指标的组合，这些指标涵盖公司的增长、盈利能力和资本回报，并提供反馈性和前瞻性观点。总体而言，我们应关注让管理团队创造长期的经济价值。[61]

毫无疑问，对金融市场的研究还在持续，也会出现评估公司业绩的新方法。但与此同时，与股份波动相比，这样的计分卡可以为管理层与股东及董事会代表之间的绩效合同提供更为相关的基础。

公司的目标是什么

如果说用股票价格变动来衡量企业价值令人生疑，那么有一个更大的问题是：公司的目标应该是什么？传统经济学家认为，企业或公司创立的目标是整个社会的经济利益。正如我们早期所看到的，一般均衡理论认为，如果管理者寻求将其个体公司的利润最大化，且消费者寻求将其通过消费获得的乐趣最大化时，那么资源将被优化地分配给社会。正如凯恩斯曾经说过的那样，“资本主义就是建立在这样一个惊人的信念的基础上的：人们出于最肮脏的动机，用着最肮脏的手段，却在某种程度上给全社会带来了益处”。因此，管理者为社会服务的最佳方式是将企业的利润最大化。[62]

我们在利润最大化中遇到的一个问题是，我们如何知道管理层实现了利润最大化呢？10%的利润是否足够好？15%呢？我们拿什么来衡量呢？传统经济学认为，衡量管理层是否出色的最佳裁判是资本市场。自利的投资者希望将资金投入能够获得最大风险回报的地方。由于债务资本获得了固定的利息，而股权收益会随公司的盈利状况的变化而变化，所以股东正是管理层做得好坏与否的最佳裁判。因此，将股东价值最大化能够实现利润最大化。

此外，许多国家都清楚地阐释了董事会和管理团队必须为股东的经济利益服务的法律义务或信托义务。美国和英国在提升股东地位方面走得最远，其他国家的法律制度，特别是许多欧洲国家和日本，也规定了管理层对股东的信托义务，但股东被视为必须满足其利益的众多利益相关者之一。利益相关者包括员工（通常由工会代表）、消费者、公司所在的社区以及政府。在荷兰，在董事会上，大公司需要由劳动者来代表，而公司的目标被定义为业务的连续性，而不是股东价值最大化。股东资本主义倡导者认为，这样的利益相关者系统在通过满足竞争性利益集团来混淆视听，管理层更不负责，实际上不如股东制度，不能较好地服务于整个社会。股东资本主义倡导者以此为证据指出，美国的股东资本主义已经创造出了高生产率、高增长和高就业的经济。[63]

这不仅仅是一个学术问题，一直以来，围绕着治理改革、企业的社会责任倡议、反全球化抗议以及欧盟经济模式的未来，人们一直争论不休，而其中最为核心和饱受热议的话题便是企业应该为谁的利益服务。[64]

持续与增长

复杂经济学并不一定与传统的论证方式相矛盾。市场仍然是定价和分配资源的有效机制。股东仍然是公司的合法所有者，管理层仍必须对其业绩负责。但复杂性理论确实提供了更细微的观点，提出了一些重要的问题，并提出了另一个目标函数。

如果回到进化框架的第一条原则，我们也能够深入讨论公司的目标问题。正如第9章中所讨论的那样，进化系统对实现目标来说具有鲜明和纯朴的逻辑——好的基因得到复制。任何模式的目标函数都必须保证其交互者能够生存并进行复制，任何其他的目标函数都会导致灭绝。在经济学当中，这意味着任何商业计划的目标都应该保证业务存活并扩展业务。管理层的职责是运用这一目标函数来运行商业计划。换句话来说，这意味着管理层的工作就是制定和执行计划，保证公司的业务能够长期持续发展。

乍一看，获得“持续和增长”的目标函数可能听起来更接近荷兰的连续性目标，而不是美国最大化股东价值的目标。但进一步的反思表明，股东价值的连续性和最大化可能并非水火不相容。

所有进化系统都在热力学定律所施加的一系列约束下运行。生存和复制需要能量、物质和信息的流动，在生物系统中，这意味着所有生物体都面临热力学定律的约束，因此随着时间的推移，生物体摄入的卡路里必须等于或大于燃烧的卡路里。优秀的复制者能够获得卡路里能量的盈利，他们会设法在资源有限的竞争环境中获利。同样，在经济系统中，熵减少也需要能量、物质和信息的流动（在现代经济中以货币计量），所有企业都面临利润限制，在这种限制下，流入必须等于或大于流出。

在进化系统中，盈利本身并不是一个目标；相反，如果企业要实现生存和复制（或持续和增长）的目标，就必须实现盈利这个基本的约束。管理思想家查尔斯·汉迪（Charles Handy）举例说明了这种差别：人们要想生存，就需要吃饭，这是我们生存的一个限制条件，也是非常重要的一个因素，但没有人会说生活的目的是吃饭。[65]

如果从经济学背景下解读盈利这个概念，我们会发现另一组约束。要成功执行商业计划，管理团队必须在组织内部和周围创建合作的生态系统。第一，必须把资金吸引到企业之中，资金的提供者必须看到比其

他投资机会更好的潜在回报。第二，必须吸引员工并激励其高效地工作。第三，供应商必须看到盈利的商业关系。第四，企业必须提供人们想要的商品和服务，但这并不是目标。管理层还必须确保企业履行其法律义务，缴纳税款，并且谨慎经营，防止普通民众想要关闭企业的情况发生。事实上，盈利能力也需要在多方面令人满意。从进化的角度来说，这就是经济适合度函数：与适合度低的公司相比，适合度高的公司能够更好地满足多维度的约束。从进化的角度来看，使股东主义与利益相关者主义对抗是一种错误的二分法。适合度函数就是这样——无论是将其称为股东主义还是利益相关者主义，这都是管理团队维系及发展业务必须要做的事。

使股东主义与利益相关者主义相对立在实践中也是一种错误的二分法。大多数股东资本主义倡导者承认，如果一家公司对员工不公正，违反法律，对待客户态度差，那也不太可能长期为股东提供较好的回报。大多数利益相关者制度的拥护者都承认公司必须获得有竞争力的资本回报，以保证投资者参与其中。在现实中，股东和利益相关者体系之间相匹配的关键因素是具有竞争力的市场，劳动力、资本、客户和供应商都必须能够在市场中流动。

传统经济学将股东与其他利益相关者区分开来另有原因，因为只有股东才能获得公司的剩余索取权（residual claims）。换句话说，所有其他索取人得到应得的一切后，股东获取剩余的一切，因此股东也承担着风险。有可能其他利益相关者将公司抢劫一空，股东两手空空，由此可见，股东承担着更多风险并需要更多保护。与合作企业相关的每个人都承担着风险——员工赌上了自己的职业生涯，而消费者面临的风险则是产品是否能够达到预期，等等。在自由和充满竞争的市场中，这些风险应该反映在价格中。股权资本也有风险，因此其比债务资本要求的回报更高。

有些人可能会争辩说，经济学当中并没有将股东置于首要位置。许多国家郑重地将股东的重要性写入了法律，尤其是美国，但这也不是真实的情况。这些法律确实规定了企业对股东的义务，但同时也规定了企业对所有其他利益相关者的义务。美国法律规定了某些利益相关方是最重要的，这里指的是美国政府。法律规定与进化经济观点没有什么不同，管理层必须履行其对所有利益相关者的义务。

从进化和实践的角度来看，向股东支付回报是种约束，而不是目标。虽然这可能是至关重要的约束，但也只是经济适合度函数中的众多

约束之一。那么，“持续与增长”的目标意味着什么？

要回答这个问题，就需要探讨企业的优越之处。哪家公司更“优秀”，是20世纪60年代和70年代股票像彗星一样崛起，在短短几年间令世人瞩目，随后又在20世纪80年代陨落的王安电脑公司（Wang Laboratories），还是股票表现在大部分时间不及大盘指数，但自1802年以来一直经营到现在，在2005年获益270亿美元并雇用6万人的杜邦公司（DuPont）？[66]又或是自公元795年以来始终在满足英格兰圣奥尔本斯市民的需求，但一直维持着乡村酒吧规模的肯德尔斗鸡老酒店（Ye Olde Fighting Cocks Pub）？

大多数人会凭直觉选择杜邦公司。虽然王安电脑公司在一段时间内迅速扩张，最终却没能维持这种状态。虽然肯德尔斗鸡老酒店的经营历史令人印象深刻，却并没有扩张。杜邦公司成功地做到了这两点。尽管其股票价格落后于市场平均水平，但随着王安电脑公司闪亮登场又销声匿迹，杜邦公司从初创的火药制造商发展成为全球主要的化学品、材料和生命科学公司，并且持续经营超过两个世纪。传统经济学家也会同意这个观点，因为随着时间的推移，杜邦公司创造的总经济价值远远大于其他两家公司。事实上，为了延续和发展，杜邦公司必须完成传统经济学家认为需要达成的每个目标。杜邦公司必须盈利，因为不盈利的公司无法延续，还必须为其资本提供者提供有竞争力的回报（即使公司本身没有获得最高回报），因为公司增长需要资本。[67]

人们可能会问：“如果说传统经济学和复杂经济学殊途同归，那么两者的差异是什么？”也许约束和目标之间的区别只是从学术用词上进行了区分。

关键的区别在于管理团队如何应用这些概念。许多管理团队十分痴迷短期股票价格波动和季度收益，已经实现了最大化股东价值的目标。一项针对美国国家经济研究局的400多名首席财务官进行的调查发现，大多数公司将季度收益视为管理的关键指标，因为他们相信管理收益能够实现公司股价最大化。[68]公布季度收益可能确实会推动股票价格上涨，但正如我们所看到的，这些消息并不是推动价格走势的唯一因素。对季度收益过于执着反过来经常会扭曲管理决策。在同一项调查中，大多数首席财务官表示，如果创造价值的长期投资缺少分析师的盈利预测，他们就会放弃此类投资，超过3/4的人表示他们会牺牲经济价值来获得平稳的收益记录。

专注于持续和增长的管理团队可能会持有更加均衡的观点。[69]团队

成员将专注于实际可以控制的杠杆，即创造经济价值的杠杆，而不是股票价格的杠杆。管理层也会更明确地认识到长期持续和增长的多维性质，虽然股东仍然是一个至关重要的支持者，但追求持续和增长的团队将会调整方向，将管理时间和注意力放在维持与利益相关者的关系上。关注公司的持续性并且将此视作目标的公司，自然会从长远角度进行投资，并且会通过强大的公司文化等因素奠定公司长寿的基础。与此同时，这些公司也同样关注企业增长，这也对公司的业绩造成了一定的压力和挑战，也会催生创新。当然，公司也可以通过投资不盈利的项目获得增长，或者投资一些回报不足的项目，这同样也会破坏经济价值。但同样，这两个目标自然会慢慢平衡——对无利可图的增长进行鲁莽的投资会危及公司的持续发展，而以公司持续发展的名义采取过于保守的策略则与增长的目标背道而驰。

以持续和增长为目标，不仅有助于管理层阐明其存在的理由并采取更均衡的决策，也有助于管理层将目标传达给重要的支持者，尤其是员工和外部利益集团。很少有员工一大早起来的时候就会元气满满地去实现股东价值最大化的目标。大多数员工没有股票期权，他们认为股东财大气粗，却不知道股东是谁。但是员工可以树立建立一个伟大而持久的公司的信念，而公司在增长的过程中也会为员工创造机会。

在企业之外，政治家和活动家也难以将股东价值最大化作为公司的目标。他们认为公司的存在是为了服务于社会的更大利益。无论如何他们都不会相信，为股东服务同时也是为社会服务。在许多管理团队关于企业的社会责任和公共关系问题的辩论中，这种思维方式造成了巨大的沟通障碍。[70]如果管理层能够向这些人表明管理的目标是建立长久运行的机构，能够为人们提供想要的产品和服务，创造就业机会，纳税，在法律范围内运作，与外部保持良好的关系，并且不断创新和发展，在未来产生更大的社会效益，那么毫无疑问，人们会更多地看到社会和企业利益之间的联系。

预料之中的是，有一些研究发现，像杜邦公司这样经营多年并且规模较大的公司通常并没有把股东价值最大化视为公司发展的目的。例如，根据著名的荷兰皇家壳牌公司（Royal Dutch/Shell）对27家大型长寿公司开展的调查研究，阿里·德赫斯（Arie de Geus）在《长寿公司》（*The Living Company*）一书中描述了这些公司如何看待“生存与蓬勃发展”的目标。[71]同样，吉姆·柯林斯和杰里·波拉斯在《基业长青》一书中研究了18家长寿公司，并将这些公司的态度概括为“不仅关注利润”。作

者补充说，对于这些公司来说，“利润就像供养身体的氧气、食物、水和血液，它们不是生命的重点，但没有它们，就没有生命”。[72]

持续和增长的目标与前一章讨论的执行和适应的必要性相呼应，这绝非巧合。两者都源于相同的进化逻辑。相比于适应调整，公司通常更擅长执行，这无疑与关注短期收益有关。管理层要欣然接受持续和增长的双重目标，就要充分地应对执行和适应中固有的紧张局势，同时他们也要在压力之下平衡两者的关系，采取更均衡的策略。在充满竞争的进化环境中，“持续和增长”是要达成的目标，而“适应和执行”则是如何达成这样的目标的路径。进化系统中，持续和增长是所有安排里永恒的需求。

18

公共政策：左右派争论的终结

经济世界总是与政治世界联系在一起，而且从历史上看，经济学理论中的范式转换导致了政治格局的重组。亚当·斯密的思想极大地促进了19世纪自由贸易的发展；马克思的理论助力了20世纪的大变革；新古典主义范式奠定了西方资本主义正统思想的智力基础；而凯恩斯主义则用国家的角色调和了西方资本主义正统思想。第二次世界大战后，西方经济学中国家激进主义思想的发展突飞猛进，直到20世纪80年代遭遇了罗纳德·里根和玛格丽特·撒切尔的挑战，而这两位都受到了米尔顿·弗里德曼和弗里德里希·哈耶克等经济学家的启发。

复杂经济学思想将如何影响政治和政策？通常经济思想的学术发展与其对社会产生的更大影响之间存在至少几十年的时间差。马克思于1848年发表了他的《共产党宣言》，然而俄国革命直到1917年才发生。凯恩斯在20世纪30年代就出版了他最具代表性的著作，而他的思想所带来的巨大影响一直到第二次世界大战后才得以体现。甚至于在《国富论》出版两个世纪以后，许多政治人物似乎仍然难以理解亚当·斯密的思想。虽然我们可能要过很多年才能感受到复杂经济学对公共领域产生的巨大波动，但不妨先来猜想一下究竟会有些什么样的影响。

关于复杂经济学与政治的关系，大多数人问的第一个问题是：它究竟站在哪一边？复杂经济学是亲左派还是亲右派？有些人可能会怀疑它是亲右派，因为我说过复杂经济学对市场持积极态度；其他人可能会怀疑它是亲左派，因为我也说过市场并没有传统正统思想所宣称的那么有效。而我的观点是复杂经济学既不亲左也不亲右。事实上，复杂经济学可能会使政治的历史框架过时。

我们先来简要回顾一下左右派框架的历史。我认为，在这一历史鸿沟的核心点上存在着两种根深蒂固的分歧：一是双方如何看待人性的根本性分歧，二是对国家与市场的角色的看法的分歧。复杂经济学为这两个由来已久的争论提供了一个全新的视角，而且它不属于左、右派中的任何一个阵营。接下来我们将讨论复杂经济学如何应用于以下三个领

域：发展中国家的贫困问题；研究人员所说的美国和其他发达国家的“社会资本”侵蚀问题以及经济不平等问题；以及复杂经济学的未来研究方向问题。

左右派时期

以左、右派格局看待政治的想法经久不衰，延续了200年之久。这些术语源自1789年法国大革命时期建立的法国国民议会的布局。处于第三等级的革命派坐在会议厅的左边，保守的第一等级保皇派坐在右边。从一开始，“左派”一词就有争取社会进步，捍卫弱势群体，改造社会使其变得更好的含义，但也带有乌托邦的味道。“右派”一词包含强调个人自由和责任、保护社会稳定及相信顺其自然、循序渐进的发展等含义，但也有保护特权和当权者的意味。到马克思、恩格斯理论时期以及社会主义和资本主义对峙时期，经济层面的左派和右派的分裂变得更为明确。[1]到了20世纪初期，左派与主张政府大力干预经济的政策联系在一起，这些政策包括主张政府从完全拥有到部分拥有共产主义经济的资产和政府在社会主义市场中扮演监管角色，而右派则是自由市场的支持者。尽管这些术语的具体含义在不断演变，但左、右派广义上的分歧在整个20世纪中后期仍然没有改变：国家与市场的对立，社会自由主义与社会保守主义的对立，以及大众需求与个人权利的对立。

1989年柏林墙倒塌后，约翰·霍普金斯大学教授弗朗西斯·福山（Francis Fukuyama）大胆预言：这场伟大的辩论终于结束，“历史的终结”已经到来。[2]在这期间，许多知识分子和政治家试图重新定义“左派”与“右派”。1996年，美国民主党领袖委员会发表了《新进步宣言》（*New Progressive Declaration*），1998年，布莱尔的顾问伦敦经济学院院长安东尼·吉登斯（Anthony Giddens）出版了一本名为《第三条道路》（*The Third Way*）的书。[3]被称为美国“新民主党”和英国“新工党”的这些人都主张将资本主义创造财富的成功之处与社会主义的人道主义目标相结合——从本质上说，就是通过资本主义的手段达到社会主义的目的。这一框架推动了克林顿和布莱尔这两个政府的经济意识形态的形成，并且在克林顿改革美国福利制度以及布莱尔试图为陷入困境的英国国家医疗服务体系注入市场力量等政策中得以体现。包裹资本主义的善意外表甚至吸引了右派人士。小布什的第一次竞选就是主要围绕着“富有同情心的保守主义”这一思想展开的。

但左、右派的对立并没有结束。西雅图街头暴民抗议1999年世界贸易组织会议的行为表明，左、右派之间的分歧确实还存在。同样，布莱尔的许多改革措施也因遭到党内左派的强烈反对而停止，小布什富有同情心的保守主义最终变成了给富人全面减税。柏林墙倒塌之后，左、右派之间的分歧可能已经有所缩小，但并没有消失。

20世纪90年代，第三条政治道路开始有所发展，它更多的是以切实可行的政治主张，而不是以新经济学理论为基础的。左、右派都吸取了教训，因为极端分子、纯粹主义者以及相关模型的运用根本不起作用。国营的乌托邦变成了官僚主义的噩梦，而恍如天堂的自由市场却导致了社会功能失调。因此，在见证了苏联解体和大萧条之后，许多人最终从实际出发达成了广泛共识，即市场和国家在社会中都有各自的作用。然而，这种中立的共识却留下了一片知识空白。尽管先辈们有这样的想法，但第三条道路已经演变成一套从中心地带赢得选举的政治策略，而不是一种真正的新经济范式。[4]

复杂经济学正好有可能填补这片知识空白，其提供的不是一个混乱的中间层——它既不是遭遇过一些市场失败的新古典主义，也不是采用了一些市场机制的社会主义，而是给出了一个全新的视角。根本性问题不是左派与右派的对立，而是如何更好地发展。

人性与强互惠

如果你深入探讨左、右派的分歧，一直深入到哲学和历史的核心，就会发现两种相互冲突的人性观。左派人士认为，人类天生就是利他主义的，贪婪和自私并非源于人性，而是源于社会秩序的构建。在一个更加公正的社会中，人类会变好，这一观点的传承源于卢梭和马克思。

右派人士认为人类天生就是利己主义的，追求自身利益是人类不可剥夺的权利。最有效的政府制度应该是包容，而不是试图改变人性的这一面。正如18世纪苏格兰哲学家休谟所说，“在设计任何政治制度时……每个人都应该被看成是无赖，人们所做的一切除了谋求个人利益，别无其他”。[5]然而，右派人士认为，人们在通过市场机制追求自身利益的同时，社会的普遍利益也会得到满足。这一观点的传承源自休谟、约翰·洛克和托马斯·霍布斯。

你可能会感到惊讶，亚当·斯密的名字居然没有出现在这里。但正如经济学家赫伯特·金蒂斯、塞缪尔·鲍尔斯、厄恩斯特·费尔以及人类学家

罗伯特·博伊德所指出的那样，亚当·斯密实际上持有一种更为微妙的观点。在《国富论》中，亚当·斯密确实写到，在市场调节下利己主义能够带来社会效益。但在他的另一部伟大著作《道德情操论》中，亚当·斯密也说过："无论人们觉得一个人多么自私，他的本性中显然还存有一些原则使他对他人的命运感兴趣。"[6]换句话说，斯密对人的行为持有一种更为全面的观点，承认人性既有利己的一面，又有利他的一面。

金蒂斯和他的同事认为，现代研究表明了历史上的左派和右派对人性的看法都太过简单化。几个世纪以来，人类的利己主义与利他主义之间的问题一直是一个哲学问题，说到底最终是一个观点问题。然而，自20世纪80年代以来，它已经成为一个科学问题。根据对照实验、实证研究、人类学的实地调查以及博弈论应用中得到的大量证据，现在我们已经找到了答案：亚当·斯密基本上是对的。

人类既不是天生的利他主义者，也不是天生的利己主义者，他们是研究人员所说的有条件合作者和利他性惩罚者。金蒂斯和他的同事将这种行为称为"强互惠行为"，并将其定义为"一种与他人合作并惩罚那些违反合作准则的人的倾向，哪怕付出的代价后续很难弥补"。[7]这就是我们先前讨论最后通牒博弈和不断演变的囚徒困境时所看到的行为。从本质上说，人们是在试图遵循黄金法则，但也有一些区别：你希望他人如何对待你，你就如何对待他人（即有条件的合作）；如果他人不这么对待你们，那就迫使他这么做，不惜付出个人代价（即利他性惩罚）。人有一种高度发达的心智，知道他们可以信任谁，不能信任谁，他们欠谁的人情，谁欠他们的人情，以及他们是否正在被利用。古话说得好："欺骗我一次，可耻的是你；愚弄我两次，可耻的是我。"

强互惠行为的普遍存在令人感到震惊，从现代工业社会的人群往回追溯到遥远的狩猎采集部落的人群，人们都发现了这一行为。对于这种行为多大程度上由遗传决定，多大程度上取决于文化，现在普遍存在着争论，但有三份证据充分指向了遗传基础。第一，有研究表明，强互惠行为出现在各种文化环境中——未发现任何社会没有展示强互惠行为的某种形式，从而表明其起源并非基于纯粹的文化。[8]第二，人们在许多灵长类物种中也观察到了类似的行为。[9]第三，科学家在催产素中发现了这种行为的生物化学基础。催产素是一种大脑激素，它在激发人类的信任感和引发合作方面发挥着重要作用。[10]虽然强互惠行为很普遍，但是它在不同社会有着不同的表现形式，它的执行方式也极其多样化，因此其发展很可能是由基因和文化共同进化决定的。

强互惠行为的进化逻辑很简单：在非零和博弈的世界中，有条件的合作者比运用纯粹利己或纯粹利他策略的主体表现得更好。尽管林格伦的囚徒困境进化模型不受单一策略的支配，但有关有条件合作者主题的顶级策略往往多种多样，这也不是一种巧合。研究人员对世界各地的人们参加最后通牒博弈所得的结果调查后发现，最接近于传统经济学所假定的利己主义合理性行为的社会团体是秘鲁热带雨林中的美洲印第安人。美洲印第安人的强互惠文化准则不如其他社会成熟，因此具有自私、互相猜疑和合作能力低下的特点。他们的社会并没有超出以家庭为单位的组织，所以他们是接受测试的群体中最贫穷的人群之一也就不奇怪了。[11]

传统经济学家可能认为，强互惠只是利己主义的另一种形式。毕竟，人们都是为了满足自己的目的才进行合作的。在非零和博弈的世界中，合作的确是有回报的，但强互惠和传统的利己主义有两个关键的区别。首先，传统经济人不在乎经济交往的过程，只注重其结果能否最大化地体现主体自身的利益。然而，实验表明现实中人们不仅关心结果，而且注重过程本身是否公平。其次，正如最后通牒博弈中所看到的，人们会惩罚不公平的行为，哪怕需要付出代价，甚至是付出后没有希望得到弥补的代价也无所谓。换句话说，当人们感觉自己真的被骗时，他们会做一些很疯狂的事情。当然，这背离了利己主义的理性。

强互惠行为对经济和政治所产生的影响可能不会立刻显现，但一旦我们改变了人类行为的核心假设，很多变化就会出现。[12]从20世纪30年代到60年代，美国政府对穷人施以援助的计划得到了广泛而又普遍的支持。而从20世纪70年代到90年代，人们对此的支持率急剧下降，而下降的原因成了许多辩论的主题。左派人士认为，支持率下降源于种族歧视，因为领取救济金的绝对是少数种族，而且在这时候，崇尚利己主义的一代崛起了。换句话说，就是缺乏利他主义者。右派人士对此的解释是，人们终于意识到大多数福利项目是无效的，只是在浪费他们的税收，现在想把交的税收回去。

克里斯蒂娜·方（Christina Fong）、鲍尔斯和金蒂斯在进行调查、实验和小组讨论后找到了重要证据，并非是左、右派的这些解释，而是强互惠行为导致了人们态度的转变。在项目创立之初，人们认为那些领取救济金的主要是些想要工作却因为运气不好、经济形势瞬息万变而无法找到工作的人。社会准则也支持这种看法，认为这样的人应该得到帮助。然而在近代，人们已经转变了观念，认为靠救济生活的人都非常懒

惰，不思进取，在滥用社会的慷慨。这样的行为违反了互惠准则，不应该对其施予援助，甚至应该对其进行惩罚。

一些人提出，社会政策应该是专门设计用来“倡导而不是冒犯互惠价值观的”。例如，符合强互惠准则的政策包括为想要工作的人提供技能培训，激励穷人储蓄，支持贫困地区的人们开展创业活动，以及增加弱势群体的受教育机会。同样，强互惠准则鼓励人们将弱势群体划分为值得援助和不值得援助两类。体现这一区别的项目往往会得到广泛支持。例如，提供失业保险的一些州立项目会很受欢迎，因为工人只要在在职期间支付保费，万一哪天他们运气不好被解雇了，就能领取失业救济金。同样，社会保险得到了民主党和共和党两党多年的支持，很大程度上也是因为它符合强互惠准则——人总会慢慢变老，那些缴纳过社会保险金的老人也可以从中受益。而违背这些准则让“不值得援助”的人受益的项目往往会引发争议，例如，发放福利的同时没有提出让受助者找工作或接受培训这样的对等要求的福利项目，或者针对吸毒成瘾的人的康复项目。但是，如果救助要求中有对等行为，例如接受福利的人必须参加工作或者吸毒者必须戒毒，那么这些项目就会得到普遍支持。

强互惠准则有助于解释美国和英国的新左派为何试图将个人责任重新纳入发展日程。克林顿推行的包含工作要求的福利改革，以及布莱尔所倡导的“严厉打击犯罪，严厉打击犯罪源头”的运动就是最好的例子。右派人士也开始运用这些准则，例如受到小布什政府支持的基于信仰的新方案就结合了社会目标与受宗教激励的责任和互惠价值观。

人类既不是卢梭笔下纯洁无私的生物，也不是休谟笔下无情自私的生物，而经济学家和道德哲学家亚当·斯密是对的——人类两者兼具。复杂经济学对强互惠准则所持的观点是，左派最终可以抛弃卢梭所认为的一切社会弊病都是社会的过错，并承认个人责任的作用。同样，右派也可以抛弃休谟的观点，即社会的建立前提是假设人都是无赖，并承认人类更慷慨的本能的作用。

复杂经济学表明，个体行为只是其中一个因素。制度的突现行为是由个体行为和制度结构一同创造的，这就把我们引向了左、右派之间的下一场大辩论——市场作用与国家作用的对立。

左派乌托邦和自由市场幻想

正如我们在第三部分看到的，人类历史上只创建了两种机制来促进

陌生人之间的大规模经济合作——等级制度和市场。[13]这两种机制中都包含了各种各样的执行方式，但最终都可以归结为一种或另一种。即使是较平等的组织形式，比如公社和合作社，也包含某些形式的权力等级结构。[14]从复杂经济学的角度看，资本主义经济和社会主义经济的关键区别在于，经济健康的最终仲裁者是市场还是等级制度。

无论是资本主义制度还是社会主义制度，等级制度内都有一个商业计划的分化、选择和扩大的过程。在资本主义经济中，这一过程发生在设置私营部门的公司中，这就是艾尔弗雷德·钱德勒所说的工作中的“看得见的手”；而在社会主义经济中，这一过程发生在由国家直接或间接控制的组织机构中。在资本主义经济中，经济进化过程最终会进入到市场的“薄弱环节”，正是在这一环节，亚当·斯密的“看不见的手”最终决定了商业计划的选择和扩大。在社会主义经济中，市场的薄弱环节并不存在（或由国家主导），政府组织内部会作出商业计划选择和扩大的关键发展决策。

右派的批判观点

复杂性对意识形态分歧另一面的批判，与其说是对资本主义本身的批判，不如说是对受新古典主义启发的右派幻想资本主义运作方式的批判。这一论点有两个关键论据。

第一，尽管复杂经济学认为市场不仅有用而且必要，但它也否定了市场的最优效率基础。正如我们前面所看到的，即使拥有如教科书市场一般的金融市场，其运作起来也远不能达到理论效率。这意味着某些右派人士认为市场是解决所有社会问题的办法这一倾向是错误的。公平地说，大多数传统经济学家都在一定程度上认识到了市场的不可靠性，而且学界也充满了对市场失灵的分析。但复杂经济学的观点将使我们越过传统的市场失灵现象，对那些依赖传统理论常用的不现实假设的市场解决方案持怀疑态度。

一个很小但非常典型的例子是英国撤销了对电话簿查询的管制。几十年来，人们在英国只能通过拨打一个特定的国家号码查询电话号码。该服务以前由英国国有电信公司垄断，但其服务价格和质量都受政府的监督和管制。这项服务简单便捷、运作良好，而且收到的消费者投诉很少。然而，2001年，政府决定解除对该服务的管制，开放竞争。理论依据是，市场竞争会使服务价格降低，让服务更具创新性。正如预测的那

样，大量新公司进入了市场，2004年该市场中有了120家公司。然而理论和现实很快就脱节了。之前只有一个号码——“192”，人人都记得。而在新市场中，人们需要面对120个相互竞争的号码，真的很少有人能记住这么多的新数字。新号码通过抽签分配，幸运的公司分配到了容易记住的号码，比如“118”，迅速成为新崛起的垄断者，占据了大部分的流量。况且，对大多数人而言，电话号码查询费用是一笔很小的开支，不值得花时间和精力去寻找价格最低的供应商。结果就是新的容易记住的号码收费更高，收到的投诉也比原来多。再者，新号码不是真正的国家号码，法规允许电话公司自行限制哪些号码可以从他们的网络接入，因此人们从家里、办公室或用手机打电话时可能需要记住不同的号码。人类认知的现实状况和制度的设计都与传统理论的简单世界相冲突，因此市场竞争的好处并没有得到体现。[15]

类似的例子还包括美国加州电力市场放松管制导致了2001年严重的电力短缺以及英国铁路网因私有化耽误了大量的火车行程。但这并不是说市场是不好的，垄断才是好的，复杂经济学也广泛支持市场是经济进化的最佳机制这一观点。相反，以传统经济学过于简单化的理论为基础的右派则倾向于将市场视为解决所有问题的方法。在每一个实例中，人类的真实理性与制度的烦琐细节都与理论不符，市场解决方法也无法为社会带来好处。

第二，在复杂经济学背景下，许多右派人士的反政府立场看起来非常幼稚。新古典理论创造了一个存在于原始自然状态下的理想经济，没有政府的干预。正是在这种理想化状态下，帕雷托最优才得以实现，并为社会创造了最大的财富。一旦税收、法规等对经济加以干涉，经济就会偏离这个理想化状态，为社会创造的财富也会随之减少。因此，他们的目标应当是将税收和政府开支控制在最低限度。同样，右派认为，虽然某些基本法规可能是确保市场运作的必要前提，但法规通常具有扭曲作用，会干扰自然价格信号，并且像税收一样让经济偏离理想状态。罗纳德·里根在1986年白宫小企业会议上的讲话中曾说过一句名言：“政府对经济的作用可以用几句简单的话来概括：动，税之；恒动，规之；止动，资之。”

但是反政府的自由放任主义者们忘记了经济并不是孤立存在的。经济进化系统是由大量的社会技术组成的，其中的许多技术依赖于政府。[16]基于市场的经济进化需要在合作与竞争之间维持微妙的平衡，而政府在保持社会的这一平衡方面发挥着至关重要的作用。诸如合同法、消费

者保护条例、工人安全法规和证券法等社会技术都有助于促进合作、增进信任，反垄断法规则有助于保持健康的竞争水平。

美国的右派人士经常指责民主党人富兰克林·罗斯福实施新政后加大了政府对经济的“干预”。但实际上将政府定位为合作与竞争之间的平衡者的关键策划者是一位共和党人，也就是另一个罗斯福——泰迪。[17]在1901年至1909年期间，泰迪进行的改革发生在从破坏洛克菲勒和摩根之间的信任到颁布第一份食品安全条例等领域，为20世纪的经济增长提供了重要的制度基础。

人们只要到政府机构软弱无力的某些发展中国家去走走，就会看到没有政府干预的生活会是什么样子。当政府不能发挥干预作用时，经济就会陷入低合作、低竞争的死胡同。并非所有的政府法规都是好的，有些规章可能是愚蠢而又多余的。然而，“政府是问题的根源，而不是解决方法”的右派言论忽视了政府机构在支持经济系统有效进化方面的重要作用。

政府作为适合度函数的塑造者

复杂经济学不仅为公司，还为政府所扮演的经济角色提供了一个新的视角。受新古典主义启发的右派认为，帕雷托最优状态确实是自由社会最好的结果。由于市场是实现帕雷托最优的最佳机制，也是唯一的机制，所以政府的正确作用是确保市场效率。任何干预市场的政府行为所发挥的作用都应当在成本效益抵消效率损失后进行衡量，比如右派常常使用市场效率这一理由来反对环境法规。

反过来，复杂经济学也对两种类型的政府行为进行了区分。让政府参与区分、选择和扩大商业计划的政策，被认为是干涉经济进化。这方面的例子包括日本出台产业补贴政策和保护受支持的产业，法国为支持本国国有银行对欧洲银行业的兼并采取干预措施，或者美国因为艾奥瓦州的政治重要性出台能源政策支持用玉米提取乙醇的项目。相比之下，在把商业机会的选择和扩大留给市场机制决定的同时，塑造适应性环境的政策则是另一回事。正如我们所看到的，进化系统的效率概念只是一时的。从复杂经济学角度看，政府法规是公司竞争所处的适应性环境的组成部分。只要市场提供了选择和扩大商业计划的机制，那么经济进化过程就会根据这些法规进行创新并适应这些法规。

如果选民告诉他们的当选代表，保护环境是社会需要优先考虑的事

项，那么政府就有责任在职权范围内调整经济适合度函数，使之有利于环境友好型商业计划，而不是非环境友好型计划。这方面的例子包括碳税、排污或强制性工业回收要求。在这些案例中，政府没有选择计划，而是塑造了计划取得成功或遭遇失败的适宜环境（例如，在一个征收碳税的世界里，低排放计划会比高排放计划更有效）。

这种将政府看作适合度函数塑造者的观点与左、右派思想格格不入。右派对效率成本相当抵触。但是，正如我们所看到的，效率成本参数本质上是静态的，例如，如果碳税激发创新使得太阳能设备成本大幅度降低，那么经济效率究竟是提高还是降低了呢？相比于市场的创造力，左派更愿意相信政府的合理性，想要采取更有指导意义的方法，例如希望对所有发电厂设定强制性减排目标，而不是发起排污交易。不过，有人可能会说，如果左派和右派都提出反对意见，那么很可能就意味着我们走上了正确的道路。

然而，鉴于哈耶克指出的所有原因，我们还是得谨慎看待将政府视为适应性函数塑造者的想法。我们应该从实际出发正视自己的预测能力，合理预测政府塑造适合度函数所带来的影响以及出现非预期后果的可能性。这并不意味着政府无法有效地发挥这一作用，相反，这意味着政府得有意愿去做试验、收集反馈以及变换方向，但所有这些在政治层面都没那么容易。它还改变了辩论方式，将市场效率和社会目标之间陈旧的意识形态斗争变成了如何调整市场进化过程来更好地满足社会需求的实践讨论。

复杂经济学不仅改变了我们对左、右派意识形态立场的看法，而且改变了我们对左、右派各自推崇的两个主体内的看法：国家和市场。国家的经济作用是建立一个制度框架，支持不断进化的市场运作，在合作和竞争之间实现有效的平衡，塑造经济适合度函数使其更好地满足社会需求。根据强互惠准则，国家有义务确保其所有公民享有同等机会参与到经济体制中，并为那些没有在这一体制中取得成功的人提供基本保障。市场的经济作用是促进商业计划的发现和其差异化的产生，将经由筛选的消费者、技术和国家塑造的适合度函数和渠道资源应用于选定的、需要扩大的商业计划。问题不在于国家与市场的对立，而在于如何将国家与市场结合起来，创造出一个有效的经济进化系统。

如果说复杂经济学还太年轻，那么它在国家政策方面的应用便是刚刚起步。并不是所有我在本书中讨论的研究人员都会明确地将自己视为复杂经济学家，我引用他们的著作是因为它们完全符合复杂经济学的一

般主题，同时举例说明复杂经济学的观点可能将我们引向何方。

贯穿这些实例的共同思路就是微观层面的行为非常重要。正如我们所讨论的，在复杂适应系统中，个体行为的规则常常会对系统的宏观性能产生深远的，有时甚至是意想不到的影响。我们探究了糖域中简单的主体规则如何导致财富的分配出现极度不平等现象；在进化的囚徒困境博弈中，主体策略如何引发主张创造性破坏的熊彼特热潮；啤酒分销游戏中的真人行为如何导致一次次的繁荣与萧条。历史上传统经济学对微观层面的行为持一种狭隘的观点：假设每个人都是完全理性的，那么探索多样化的个体行为如何影响宏观层面的结果就变得毫无意义。不过我们会发现，某些因素，比如文化准则、个人对于居住地和电视节目的选择以及父母的行为，可能就是产生重要宏观经济问题的根源。

文化的重要作用

世界银行的资料显示，2002年世界上最贫穷的国家是刚果民主共和国，人均收入约为100美元。[18]这个数值只略微高于布拉德福德·德隆估算出的大约15 000年前狩猎采集部落92美元的人均收入。与此同时，世界上最富有的国家卢森堡的人均收入为39 470美元，或者更直接地说，一个卢森堡人的年收入相当于394个刚果人的年收入之和。然而，世界上的贫富差距远远超过刚果和卢森堡之间的差距。目前，世界上有21%的人每天靠不到一美元的收入生存，其中包括撒哈拉以南几乎一半的非洲人和1/3的印度人及东南亚人。[19]正如经济史学家大卫·兰德斯所说，“世界上的国家大致可分为三类：国民花大量钱减肥的国家；国民为生存谋食的国家；国民吃了这顿没下顿的国家”。[20]

历史上对产生这种差距的解释分成了左派和右派两个阵营。[21]左派最喜欢的解释包括殖民主义、种族主义、资本主义剥削以及富裕国家给予的援助不够。而右派的观点集中在政府不作为、腐败、缺乏自由市场、依赖外国援助，甚至是微妙的种族劣势。此外，还有一些不太带有政治色彩的解释，其中包括地理环境、气候以及在非洲尤为典型的战事连连的悲剧。[22]

1999年，哈佛大学国际研究中心教授劳伦斯·哈里森（Lawrence Harrison）和他的同事塞缪尔·亨廷顿（Samuel Huntington，他对冷战结束后出现的文明冲突所做的预言非常出名）召开了一次题为“文化价值和人类进步”的研讨会。该专题研讨会将世界各地的研究人员几十年来所

做的有关文化在经济发展中所扮演的角色方面的工作结合在一起。出席这次研讨会的成员包括著名的西方经济学家、历史学家和社会思想家们，如大卫·兰德斯、迈克尔·波特、杰弗里·萨克斯（Jeffrey Sachs）、弗朗西斯·福山和内森·格雷泽（Nathan Glazer）。这次会议不仅包含西方学者对遥远文化的抽象讨论，还包括一些来自被讨论国家的研究人员和社会评论家，以及来自慈善组织的从业人员。该研讨会的成果之一是出版了一本名为《文化的重要作用》（*Culture Matters*）的书，书名正好总结了他们的讨论结果。[23]

在第16章中，我们将文化定义为受个体行为的微观准则所支配的即时产物，并讨论了文化在各类组织的经济效益中所发挥的作用。哈佛大学研讨会的与会者都认为，文化在国家经济发展中也发挥着重要作用。对于国家而言，行为规范或行为准则不仅仅是成千上万的人在遵循，而是有千百万人在遵循。

有关文化与宏观经济效益之间的联系的评论至少可以追溯到上一个世纪之交，追溯到德国社会学家马克斯·韦伯的著作。但在20世纪50年代和60年代，经济成果这样的文化解释失宠了，原因有两个。第一是政治正确性。正如兰德斯所说，“文化……让学者们感到害怕。它自带一股种族和遗传特征的酸臭味，那是一种无法改变的气味”。[24]第二个原因是新古典主义经济学占据了主导地位。在一个完全理性的世界中，从某种程度上看，文化确实没什么地位，文化准则必须是符合自身利益的优化策略，否则人们就不会遵循。[25]

大家对政治正确性的担忧已经渐渐消失，因为学者们表示完全有可能对既拥有科学成果又尊重人类多样性的文化进行讨论。你必须避开相对主义的陷阱，不要回避为什么某种文化准则可能比另一种文化准则更有利于经济发展这一问题的解释，但与此同时，你会意识到取得经济上的成功无法依靠单一的文化模式。在像日本和挪威这样文化多样而且属于经济成功型文化的世界里，主张单一完美模式的观点很容易就被否定了。同样，行为经济学的兴起逐渐减少了新古典主义假设所产生的影响，使文化这匹马在经济的马厩中站稳脚跟。

于是，至关重要的问题便成了哪些准则有利于经济发展，哪些准则不利于经济发展呢？这一领域要做的研究还有很多，但各类研究人员，比如阿根廷学者马里亚诺·格隆多纳（Mariano Grondona），已经提出了文化准则类型学。[26]这些类型学中的准则的大致分类与第16章中提到的组织规则一样，分为三大类。这也并不奇怪，因为如果组织和社会的

准则想要有利于财富的创造，就得支持经济发展进程。

第一类是与个体行为相关的准则。其中包括职业道德、个体责任，以及相信自己才是生活的主角，命运并不是由随心所欲的上帝或大人物掌控的信念。宿命论会大大减弱个人动机，重要的是我们得相信，努力工作和过有道德的生活不用等到下一世，现世就会获得回报。经济成功型文化似乎在认为有可能改进的乐观主义与如实看待现状的现实主义之间达成了平衡。

第二类是与合作行为相关的准则。最重要的信念就是，生活是一种非零和博弈，合作是有所回报的。一个相信财富蛋糕固定的社会很难达成合作，而且社会成员之间往往缺乏信任。与我们所讨论的强互惠准则一致的是，文化中既要有奖励慷慨与公正行为的准则，也要有制裁那些占人便宜、欺骗他人的人的行为准则，这一点非常重要。

第三类是与创新相关的准则。[27]如果演绎部分铿锵有力，那演绎修正工作就会更加有效，因此，注重对世界进行理性、科学的解释，而不相信宗教或神秘解释的文化往往更具创新性。同样，文化需要容忍异端邪说和实验，因为严格的正统观念会扼杀创新。重要的是文化需要有利于竞争，能够歌颂成就，因为太过平等的文化会削弱人们冒险的动机。

第四类准则对前三大类准则而言都很重要，即人们如何看待时间。今天的文化（或者说是停留在过去的文化）在各个方面都存在问题，包括从低水平的职业道德到无力开展复杂的合作和低水平的创新投资。如果明天不重要，那为什么还要努力工作，为什么要进行投资合作和创新呢？相反，具有未来投资理念的文化往往看中工作，有很高的代理储蓄率，愿意牺牲短期快乐来获取长期收益，并且合作水平很高。

就像组织一样，那些拥有吃苦耐劳、合作和创新文化精神的社会，显然有利于发现、执行和进一步发展复杂的商业计划。同样，任何文化领域的严重缺陷都有可能减缓或阻碍经济的发展。

这些通用的准则无法通过单一的方式来实施。日本社会鼓励稳定合作的方法就与挪威有很大的不同，因为任何一种文化都不可能适用于所有领域。此外，对这些准则所作出的判断并不一定都是道德层面的。其中一些准则具有经济和道德的双重含义。例如，制裁欺骗行为就对经济和道德都有好处。但是也有许多准则并不具备道德内涵。例如，一个崇尚享受当下的社会和一个追求投资未来的社会从道德层面看并没有差别，但前者往往不如后者富裕。尽管我们的重点是取得经济上的成功，

但这并不是衡量社会健康与否的唯一标准。例如，一个社会可能有强有力的容忍和宽恕准则，允许欺骗行为大肆扩散开来，削弱其建立大规模稳定合作的能力，从而降低经济成果，但同样的准则也可能使它成为一个温暖、友爱、和谐的社会。

记住这些后，你就可以开始分析各个社会的文化并评估其准则所带来的经济效益。哈佛大学研讨会的一位与会者、非洲经济领导工作者丹尼尔·伊东加-曼盖拉（Daniel Etounga-Mangeulle）指出，尽管非洲文化千差万别，但“共同的价值观、态度和制度的基础，使得撒哈拉以南的各个国家密切联系在一起”。[28]他认为太多的通用准则都站在了文化类型学错误的一面，他还强调了导致非洲文化对经济产生消极影响的两个因素：一个是权力过分集中在大人物手里，另一个是看重过去和现在而非未来的时间观：“没有对未来的动态感知，就没有规划、没有远见、没有情景假定，没有策略去影响事件的进程。”

这些准则对人类产生的影响不仅仅是出现累加效应和线性效应。主体遵循文化准则进行互动，从而创造了复杂动力学。举个例子，我们来看看那些相信世界是零和博弈的人和那些把世界看成是非零和博弈的人之间的互动。如果你偏向于将这个世界看作零和博弈，那么你的目标就是得到自己的那一份蛋糕。你会将别人的收获看作你的损失，而且你的合作倾向也会很低。这样的人会投入精力去探寻更大份额的现有财富，而不是去寻找新的、更复杂的、可以创造财富的合作性活动。不难想象，在这样一个崇尚零和博弈的社会里，偷窃、欺诈和腐败现象出现的概率会更高，而且大家对这类行为所持的道德态度也会有所不同，比如偷窃行为可能会被解释为小偷只是从财富超额的人手里“拿回属于自己的那部分”。

假设一群人中有部分人认为经济蛋糕的大小是固定的，而另一部分人则认为经济是非零和博弈的。随着时间的推移，当主张非零和博弈的人找到合作方法创造出新财富的时候，主张零和博弈的人就会对其产生攻击，试图抢夺他们的财富份额。这种冲突会减少合作带来的回报，最终主张非零和博弈的人会觉得合作不划算，变成一个零和博弈的支持者。主张非零和博弈的人的态度并不是一种源自遗传的固定特性。同一个社会中，即便人们与生俱来的素质和态度完全相同，还是会出现观点对立的两个群体。但在合作力低下的社会中，人们对非零和博弈主张会随着时间的推移被消磨殆尽，最终变成主张零和博弈的人。

研究人员在建立这些动态模型的时候通常会发现一个临界点：一旦

某个群体中的不合作者和合作者人数之比超过阈值，就很难维持大规模的合作，从而落入“贫困陷阱”。[29]这样的临界点意味着，历史的瞬息万变可能会将社会沿着低合作之路推入贫困陷阱，也可能将另一个拥有着同样合作倾向的社会引领上致富之路。因此，合作者和叛离者之间的动态互动会影响社会准则的进化以及整个社会的信任水平。文化不是一股一成不变的力量，相反，社会中的人在进行互动的同时文化也在进化——文化是历史的产物，而历史又是文化的产物。[30]

大量研究表明，不同文化中人与人之间的信任度千差万别。[31] 1996年，一项重大调查在许多国家同时展开，调查人员以母语询问大家：“一般来说，在跟别人打交道时，你觉得大多数人是可以相信的还是再怎么小心都不为过？”[32]大家的回答多种多样，着实惊人。挪威和瑞士分别有65%和60%的受调查者认为他们可以相信别人，是信任度最高的国家。而信任度最低的国家是秘鲁和巴西，分别只有5%和3%的人认为他们可以相信别人。

信任和经济成功之间存在着重要联系。高度的信任感会促进经济合作，带来繁荣，反过来进一步增强人们之间的信任度，进入一个良性循环。不过这种循环也可能是恶性的，信任度低下会导致合作力低下，从而导致贫困，进而削弱成员之间的信任度。这种因果关系并不是绝对的，因为信任度并不是决定合作水平的唯一因素。举个例子，印度就认为自己国家的人民之间的信任度比美国要高。一种解释是尽管美国人之间的信任度很低，但因为他们有先进的技术，特别是有尊重法治的传统，所以美国人能够取得经济上的成功。另一种较为悲观的解释是美国人正在依赖过去的社会资本生活，除非社会资本重建，不然信任度会持续遭到削弱。有种看法是，美国民众信任度的削弱步伐最终将赶上美国经济发展的步伐，具体我们将在下文中进行剖析。

弗朗西斯·福山对为何某些文化形成的社会信任度高，而另一些文化形成的社会信任度低这一问题进行了调查。他得出的其中一个有些荒谬的结论是：在拥有很强的“家庭观念”的社会中，人们很难建立广泛的社会信任。[33]福山认为，以家庭为中心的文化虽然在家庭内部建立了很强的信任纽带，但家庭以外的可信任范围往往较小。这类文化中，血缘关系是经济网络的基础，经济网络的发展规模自然就受到了限制。福山拿那些在中国、韩国和意大利占比很大的小型家族企业及跟自身大规模经济形成鲜明对比的一小部分全球企业进行了比较。同时，那些传统家庭观念较弱的国家，比如日本、德国、荷兰就发展出了一系列社会技术来

保护家庭外部的这种信任关系。这些社会技术包括早期建立的法律法规、自发性组织和社会组织。因此，与其他文化相比，尽管这种文化中家庭内部的信任纽带较弱，但人们相互信任的范围却更广。比如，像荷兰这么小的一个国家竟然拥有荷兰皇家壳牌、飞利浦及荷兰银行这样世界级的大公司，简直太了不起了。

来自世界银行的卡拉·霍夫（Korla Hoff）和印度统计研究所的阿里杰特·森（Arijit Sen）也都认为，家庭关系太过紧密会对经济发展产生负面影响。[34]在北欧新教传统中，家庭往往被定义为由父母和孩子组成的直系核心家庭，跟爷爷、奶奶、外公、外婆、叔叔、阿姨和表亲的联系比较弱。在许多非洲和南亚国家中，家庭的范围更大，人们的家庭关系更密切，延伸到了爷爷、奶奶、外公、外婆、叔叔、阿姨、叔祖父、叔祖母、表亲、远房表亲等。在崇尚这种数代同堂大家庭的社会中，家庭成员之间分享经济财富也有着非常严格的准则，富有的家庭成员被认为理应帮助贫困的家庭成员。虽然数代同堂大家庭的成员之间互相分享、其乐融融的景象听起来很吸引人，同时还能带来心理或其他方面的好处，但从经济角度来看，它制造了一个基本问题：你找到工作，为之努力，攒了一些积蓄，结果你那没用又懒惰的远房表亲来投靠你，他吃你的，住你的，靠你辛苦挣来的工资生活，就形成了经济学家所说的道德危机。大家庭的定义引发了搭便车行为，从而降低了工作和积蓄所取得的回报。霍夫和森认为，大家庭不仅会在个人层面制造问题，当它的裙带关系延伸到商业和政府机构时，还会进一步阻碍经济发展。

尽管人类才刚开始探索文化在发展方面的意义，但越发清晰的一个事实是，文化必定是发展等式的一部分。世界银行的经济学家威廉·伊斯特利指出，从1950年到1995年，发达国家向发展中国家提供了超过1万亿美元的经济援助，然而，绝大多数受援国的贫困率几乎没有变化，自20世纪60年代独立以来，许多非洲国家甚至出现了经济倒退的现象。[35]虽然帮助减轻当下严重的贫困现状至关重要，但是人们必须更多地去了解和解决根源问题。尽管文化不是唯一的根源问题，但如果继续忽视贫穷的文化基础，那么任何援助方案都注定会失败。

社会资本与“大破坏”

2000年，哈佛大学公共政策教授罗伯特·帕特南（Robert Putnam）出版了一本名为《独自打保龄球》（*Bowling Alone*）的书，将“社会资

本”这个词带入公众视角，激发了经济学家和政策制定者对这个话题的兴趣。[36]帕特南将社会资本定义为“个人之间的联系，及由此产生的社会网络和互信互惠准则”。[37]用复杂经济学的术语来说，如果文化准则为主体行为提供了微观规则，那么社会资本就是主体创建合作网络的即时产物。

并非所有的合作性活动都能创造社会资本。例如，市场交易中的匿名合作并不能创造社会资本。拥有社会资本网络的一个重要特征是，其中的个体不断地进行互动——创造社会资本是一项接触性运动。比较典型的拥有社会资本网络的例子包括地方社区组织、慈善机构、宗教组织、运动队、俱乐部、民间团体、业余组织，当然还有保龄球联盟。公司和其他工作场所也是社会资本的重要来源，因为人们除睡觉以外的大部分时间都在工作，因此会与同事建立重要的个人关系网。

帕特南指出，虽然社会资本听起来很友好，但它并非总是好的。有组织的犯罪团伙和恐怖组织成员之间也存在高度的信任、极为有效的关系网络和强有力的行为准则。[38]为防止有人给美国联邦调查局或中央情报局通风报信，这些暴力团伙的信任纽带往往局限于至亲、宗教或族裔群体成员，以及其他知情的同伙。不过也正因为这样，团伙的发展才受到了限制。

20世纪70年代以来，帕特南一直在研究社会资本与经济效益之间的关系。作为研究的一部分，他的调查深入到了富庶的意大利北部和贫穷的意大利南部。在《让民主运转起来》（*Making Democracy Work*）一书的引言中，帕特南描述了他前往南部普利亚大区政府办公室参观的场景：

> 昏暗的接待室里躺着几个懒散的官员，他们可能每天也就出现一到两个小时，但即便在这一两个小时内也不作为。稍微多停留一会儿的访客可能还会发现，稍远处的办公室里只有一排排空空如也的桌子。一位市长因无法获得该地区官员的支持万分沮丧，对着我们大发雷霆：“他们不回邮件，甚至都不接电话！”[39]

帕特南拿这种氛围与北部艾米利亚-罗马涅地区的政府办公室氛围做了对比：

> 地区总部采用了玻璃外墙，感觉就像是一家现代化的高科技公

> 司。热情的前台人员会毕恭毕敬地将你引到相应的办公室，那儿的官员可能还会给你调出与该地区存在的问题和地区政策相关的数据库……作为许多领域的立法先锋，艾米利亚政府从言辞到行动都取得了很大的进步，它的成效通过该区域的几十个日托中心、工业园区、话剧团、职业培训点就可以看出来。[40]

其实从普利亚到艾米利亚-罗马涅也就600多公里，约6个小时的车程。两个地区居住的都是意大利人，历史背景相同，但两个政府的办事效率怎么会有如此巨大的差距呢？

帕特南把这两个地区之间产生巨大差异的根源归结为我们非常熟悉的一个话题：信任度低导致无法实现大规模合作。帕特南认为，信任程度的差异源于这两个地区的不同历史，可以追溯到中世纪时期，这一历史导致了两种截然不同的社会资本的传承。南部地区历史上实行的是君主制，等级森严、封闭、由教会控制。这样的等级结构认为民间组织和商业网络是对其权力的威胁，必须阻止它们的发展。而北部地区更加强调平等、共有、贸易自由，后来还受到了启蒙运动思想的影响。在这样的环境下，社会网络不断发展和繁荣。[41]正如帕特南所指出的，一个国家所走的道路决定了它的历史。几个世纪以来，北部的公有制共和国逐渐创造了大量的社会资本，而南方的君主制共和国却从未实现过社会资本发展的良性循环。从长远来看，这种政治上的分裂转化成了两种截然不同的经济结果：意大利北部的某些地区成了整个欧盟最富有、发展速度最快的地区，而南部的某些地区却是整个欧盟最贫穷的地区。

在《独自打保龄球》一书中，帕特南把视角转向美国去研究这个问题："美国的社会资本现状是怎么样的？"答案是"并不好"。书中开篇就写到宾夕法尼亚州经营了半个世纪的桥牌俱乐部消失了，弗吉尼亚州那些有着百年历史的全国有色人种协会、海外战争退伍军人协会、慈善联盟、缝纫妇女协会和仪仗队统统都解散了。[42]所有这些传闻也通过数据得到了证实。美国政府的可信度下降了一半，从1966年40%的巅峰值降到了2003年的20%。同样，商业可信度也从55%降到了16%，宗教机构的可信度也降了一半，降到了23%。[43]更令人担心的是，个人之间的可信度也在逐渐下降。1968年，美国50%以上的人认同"大多数人是可信的"这一说法，但到2003年这个比例降到了30%以上。信任度下滑的意义远没有我们所观察到的"在信任度更高的社会中居住会更快乐"那么简单。正如前文所说，的确有证据表明社会信任和宏观经济效益之间存在重要联系。

弗朗西斯·福山对帕特南的研究做了补充，他进一步观察到，在20世纪60年代，标志着美国社会资本下降的几个指标开始同时出现变动，包括1963年到1993年暴力犯罪率增长了7倍（近几年有所下降），同一时期的离婚率和单亲母亲生育率也出现了大幅度上升。[44]福山把1960年至2000年这一时间段叫做“大破坏”时期，并指出社会资本的下降并非只在美国出现，许多其他发达国家也出现了这一现象。

福山和帕特南都认为造成社会资本下降的原因很复杂，不能归咎于单一因素。而且双方都指出，许多因素都是双刃剑。就其本身而言，这些因素对人们有积极意义，但会对社会产生消极的影响。例如，福山和帕特南都举例说明了这一时期家庭结构发生的巨大变化，尤其是离婚率上升，女性开始进入职场。毫无疑问，离婚程序简化使得数百万人摆脱了不幸的婚姻，但不可否认的是，这也对孩子、对人的社会化以及对他们自己的信任纽带产生了一定的影响。

向女性开放职场在让女性发挥其个人潜力方面有着深远的影响。然而，正如帕特南所说，女性历来在创建社会结构方面扮演的角色都比男性更为重要。女性通常会更积极地参加社会组织、基层政治和宗教组织，而且历来都是女性在打理与朋友及邻里的关系。帕特南的一项研究表明，年龄、社会经济地位和婚姻状况相同的女性中，参加工作的女性参加志愿活动的比例下降了50%，朋友之间的日常往来下降了25%，俱乐部及教会活动出席率下降了15%，参加娱乐活动的比例也下降了10%。[45]帕特南并不是苛责女性选择出去工作，从个人层面看，女性出去工作既是经济上的需要，也是女性实现个人抱负的方式。相反，他只是指出，这一微观层面的变化——较短时间内有数百万妇女加入了工作，对整个社会经济产生了宏观的影响。

社会布局的变化，特别是郊区的扩大，也对信任和社交网络的建立产生了多方面的影响。首先，街坊邻居有了更多的私人空间，几乎没什么机会碰面、交谈。其次，人们选择开车而不是走路进出社区。这就意味着大家单独待在车里的时间久了，更没什么机会互相碰面了。而且，家与单位离得远，上下班通勤所需时间更长，从而导致职场的社交机会也减少了。（“我下班后也想聚聚，但是我得花一个小时才能回家见到我的家人。”）20世纪50年代，像欧吉和哈丽雅特这样的同事就住同一条街，她们下班后还会顺便一起去喝杯马蒂尼，而这在今天已经很少见了。

产生这一重大后果的另一个因素是大众媒体的隔离效应。当代人的

娱乐方式从舞厅、音乐会、酒吧、桥牌俱乐部、保龄球联盟等团体活动戏剧性地转变为个体活动。收音机、电视机、立体音响、视频游戏、家庭影院以及互联网的盛行，意味着大家留出了更多时间在家独处或者陪伴家人。而与媒体本身同样重要的是信息，媒体行业的经济现状使得从业人员将内容指向了大众，因此节目反映的往往是人类情感的最低共同标准。有些人认为，这是随着互联网内容越来越多慢慢发生变化的，但鉴于色情和网络赌博一直在流行，这样的说法还是值得质疑的。几十年来，不仅是小说中，新闻媒体中也一样，整个社会充斥着谋杀和蓄意伤害。人类一直喜欢有关性和暴力的故事，但是通过远程控制或者点击鼠标就能把它们集中起来，让它们生动地展现在电话、电影及电脑游戏中还是人类历史上的一种新现象。第6章中，我们讨论了人类决策方面的偏差，即我们在数据推理方面有限的技巧和根据非典型样本得出结论的倾向。如果某个人置身于源源不断的谋杀新闻中，他就会下意识地认为谋杀是件很常见的事情，而且会高估自己被谋杀的概率。媒体造成的这种恐惧进一步削弱了信任，使人群变得孤立。

同女性进入职场一样，人们选择住在郊区、选择开车、选择看电视的理由也非常充分。于是，这些个体选择所产生的宏观效应又让社会资本出现了令人担忧、始料未及的下滑。社会资本长期下滑对美国经济效益所产生的影响目前尚不明确，但可能并不乐观。

鉴于引起破坏的原因的结构性和几乎不可逆转的特性，例如几乎没什么人会主张女性回归家庭，大家也不会放弃汽车，我们能做什么确实成了难题。帕特南提出了“社会资本家的议程”这一概念，其中包括以下五个部分：[46]

1．加强个人参与社会活动及少看电视的承诺；

2．在学校实施帮助下一代建立信任和社会资本准则的方案；

3．进行改革，使工作场所更有利于人们兼顾家庭；

4．改善公共交通，改写区域规划法；

5．鼓励投票，提倡政治参与。

帕特南承认，通过多方面的共同努力来重建美国社会资本是个巨大的挑战，而且我们的政治机构恐怕也无法领导这种变革工作。

福山的态度更为乐观，他认为这既不是人类面临的第一次大破坏，

也不会是最后一次。[47]他从进化论的角度指出，人类社会系统具有很强的适应性。在过去的社会大破坏时期，社会资本也出现了数十年的下滑，但紧接着的就是重新制定规范，重建社会资本。例如，福山列举了19世纪初人类从农业经济向工业经济转变所引起的巨大破坏，以及随之而来的轻微犯罪、暴力、酗酒、未婚生育以及其他社会问题增加的迹象。他指出，在维多利亚时期，针对这些变化产生的是新的社会秩序和规范，而且维多利亚时期被认为是社会资本非常高的一个时期。

围绕社会资本和信任产生的这些问题远远超出了经济学领域，实际上涉及社会的各个方面。[48]这些问题也证实了复杂适应系统是在不断变化的。传统经济学领域不可能拥有社会资本，因为传统经济学框架关注的是理性个体追求自我利益的选择，只要每个人为自己作出最优选择，就是在为社会作出最优选择。复杂经济学对此持不同观点，其认为一个社会中复杂的互动意味着善意的个体选择有时候可能带来令人吃惊的结果。在这种情况下，个体关于居住地、是否外出工作或者如何度过闲暇时光等这些不断变化、看似毫无关联的因素的选择会导致社会凝聚力急剧下降。

虽然我同意福山的进化论观点，也相信我们会找到方法去适应，但进化论无法保证一定会实现更好的社会秩序，取得更多的财富。进化过程中社会可能会经历崩溃、进入死胡同、陷入贫困。[49]因此，我同意帕特南的观点，即必须采取一种更为积极的方法，一种将个体行为、学校、政府、民间组织以及媒体结合一体的方法。

传统的左、右派理论框架很难对社会资本相关问题进行分类。一方面，合作是经济繁荣的核心这一主要观点更符合左派的思想，并且与右派追求个人主义果实的强烈思想格格不入。左派人士认为，社会资本下降会带来西方世界市场经济中已经出现的一个问题，这个问题单凭市场根本无法解决。但是在这一政治立场的对立面，那些关注社会资本下降的人则倾向于支持回归传统价值观——信任邻居、参加志愿服务、尽公民义务、参与宗教活动以及拥有稳固的家庭关系，长期以来这些一直都是保守的右派思想的主要内容。这一争论并没有把我们带入既要有社区又要有传统价值观的尴尬局面，而是两者齐头并进——拥有许多社区、崇尚许多传统价值观，以及政府的领导角色得到充分发挥。

不平等、社会流动性和贫困文化

我们已经看到了贫穷、文化和社会资本之间的关系，这里再讨论一个相关的问题：不平等。正如前文中所提到的，世界上有21%的人口，相当于有10多亿人每天的生活费不到1美元。尽管像印度这样的国家已经取得了一定的进步，但许多经济学家还是会说各国之间的不平等仍在加剧。[50]

毫无疑问，经济的出现必然伴随着经济不平等现象。一位名叫阿布尔马吉的数学家对考古数据进行研究，发现了古代埃及社会财富的分配情况。[51]他以公元前1370年至公元前1340年间人们居住的房屋面积大小作为财富衡量指标，推断出财富的分配是由帕雷托分布决定的，其参数与现代社会的参数相似。

不平等现象合乎道德吗

尽管不平等现象持久而又普遍，但它合乎道德吗？政治评论员马特·米勒（Matt Miller）在他的《百分之二的解决方案》（*Two Percent Solution*）一书中写到，如何回答这一问题是区分左派和右派的关键。[52]

右派人士说，如果生活中人们可以自由作出选择，而且市场也充满竞争，那么这种由自由人类和颇具竞争性的市场导致的不平等现象完全是合乎道德的。就像米勒所说，市场会在道德层面奖励高尚的行为，比如工作、负责任、勤俭节约、创新、承担风险。大家可以对不幸之人表以同情和怜悯，但我们没有更好的办法来分配资源。如果一个人认为市场是组织经济活动的最佳方式，那么也就必须接受其结果。

左派人士认为市场是人类为达到自己的目的而创造的、以社会技术为基础的社会构建。如果市场经济是一个社会构建系统，那么我们必须对其结果负责。在那样的系统中，并不是单凭美德就能决定结果的。一个人出身的好坏起到了很大的作用，尤其是可以遗传的智力、长相、种族，传承的财富和出生的地理位置。鉴于这些因素的影响，像米勒所说的那样去思考市场结果是否像假定的那样合乎道德没有任何意义。

这两个对立的观点还可以用强互惠准则来解释。简单地说，如果一个人的贫富是由其自身行为决定的（右派观点），那么每个人所得到的便是应得的。但是如果一个人的贫富很大程度上靠的是运气或者不受个人控制的外在因素的影响（左派的观点），那么他们得到的就不是应得的。

对于利己主义和利他主义的问题，很长一段时间以来，大家都认为

观点不同只是因为意见不一样而已。多亏许多研究人员孜孜不倦的工作，目前已经搜集到了与这一问题相关的大量证据。

近来，圣塔菲研究所的塞缪尔·鲍尔斯、马萨诸塞大学的赫伯特·金蒂斯和陶森大学的梅丽莎·奥斯本·格罗夫斯（Melissa Osborne Groves）对此进行了调查，结论是左派和右派的观点都并不完全正确。[53]

缺乏社会流动性

右派观点正确的前提是社会流动性很高。霍雷肖·阿尔杰（Horatio Algers）笔下的孩子们凭借强大勇气和毅力通过勤奋努力地工作最终由贫穷走向了富裕。这样的故事的确存在，尤其是在美国这样的国家，但实际上整个美国社会的流动性却比大家预想的要低很多。汤姆·赫茨（Tom Hertz）做了一项研究，研究中他引用了对6 273组家庭进行的长达32年以上的观察所得的数据，研究表明，非常贫困和非常富裕的地区的人群流动性都特别低（见表18-1）。

表18-1　父母的收入水平对孩子成年期收入水平的影响

鉴于父母的收入水平，儿童在成年时达到特定收入水平的可能性			
	父母的收入水平		
	高收入水平	中收入水平	低收入水平
孩子成年期的收入水平			
前 10%	29.6	7.3	1.3
前 30%	59.1	23.0	9.3
中间的 40%	34.3	48.0	28.3
更低的 30%	6.6	29.1	62.4
后 10%	1.5	7.3	31.5

资料来源：数据源自汤姆·赫茨的“黑人和白人家庭的流动性”研究。

如果一对父母位列收入最高的1/10人群，那么他们的孩子有29.6%的可能性也会成为收入前1/10的人。再往下看表格中收入在前1/10的人群那一列，大家可以看到他们的孩子有59.1%的可能性成为上流社会的人（前30%），而且他们的孩子成为低收入人群的可能性微乎其微。相反，如果父母的收入属于最低的1/10，他们的孩子有31.5%的可能属于收入最低的1/10，有2/3可能仍旧处于较低的社会阶层，只有1.3%的人可能像霍雷肖·阿尔杰笔下的孩子一样进入上流社会。而那些父母属于中产阶级（中间的1/10）的孩子很可能仍旧是中产阶级。缺乏流动性的现象在黑人当中尤其显著，出生在收入最低阶层的白人中只有17%的人会

停留在那个阶层，而黑人的这一比例达到了42%。

“哎！”左派人士为此哭诉说，“这就证明我们生活在一个不公平的社会，这个社会设置了障碍阻挠穷人和少数群体取得成功。”但是不要这么快下结论，缺乏流动性的原因多种多样，我们得跳出大众视野看看这究竟是怎么回事。[54]

我们需要考虑父母给孩子多少钱这个因素所产生的影响。这一因素解释了父母收入与孩子收入之间存在大约12%的相关性，但无法解释子女从这种给钱多少的差异中究竟获得了什么。

接下来我们来看看一直备受争议的先天遗传与后天培养的问题。也许高收入的父母会把这种高收入基因遗传给子女。对异卵双胞胎的研究表明，父母和子女的收入之间有12%的关联性可以由基因来解释，28%受环境影响。而另一项研究的结果表明，收入跟智商的关联性为5%，这就表明可以由基因解释的12%的关联性中，有其他基因特性在起作用，而不仅仅是因为天资（以智商来衡量）。因此，尽管基因是个非常重要的因素，但并不是决定性因素。

其他研究关注了诸如教育和种族这样的环境因素。贫穷的父母很可能把孩子送到贫困的学校就读，种族歧视也可能是导致父母和孩子都贫穷的一个因素。令人感到惊讶的是，学校教育对父母与孩子收入相关性的影响只占到10%，这就表明肯定有其他环境因素在起作用。在控制其他人口因素的情况下，种族对父母与孩子收入相关性产生的影响只占到7%，所以黑人与白人之间的差异肯定比我们看到的要多。

既然父母与孩子的收入之间有很大的相关性，那就是说财富实际上是可继承的。但明显能从父母那继承而来的两个因素——金钱和基因所产生的影响只占关联性的一小部分。与此同时，贫困学校和种族歧视等环境因素似乎并不能填补这种解释漏洞。这些不同因素的影响只占总相关性的1/3左右，很明显我们还没找到孩子从父母那儿获得的更关键的东西。许多研究人员认为等式中所缺失的是由文化驱动的行为和社会资本，也就是前面两部分的主题。

贫困文化

在一篇题为《苹果不会落在离苹果树很远的地方》（*The Apple Does Not Fall Far From the Tree*）的文章中，一个五人研究小组对父母

与子女行为之间的相关性进行了研究。[55]研究小组将这些行为分为“亲社会行为”，比如完成学业、去教堂、参加学校俱乐部；“不良行为”，比如吸毒、15岁前发生性行为、斗殴、破坏财产、休学等。研究人员还研究了一些会产生经济后果的个性特征，比如自负、抑郁和羞怯。研究人员发现，正如他们的论文标题所暗示的那样，父母与孩子在这些行为和个性特征上有显著的关联性。完成学业、去教堂做礼拜以及参加俱乐部的人的孩子很可能是正人君子，而爱打架斗殴者或吸毒者的孩子很可能会有不良行为。有趣的是，研究人员并没有发现社会经济地位对此有太多影响。令人惊讶的是，研究人员从参与、监督、儿童自主性培养、情感陪伴和认知刺激五个方面进行衡量后发现，父母的教养方式对孩子的影响不如预期的那么重要。真正重要的似乎是父母自身的行为表现，而不是教育技巧本身，父母的那句格言“照我说的做，不要照我做的做”似乎并不起作用。

研究人员得出的结论是，父母的行为会对孩子的行为产生极大的影响（对于这一点任何父母都不会觉得奇怪），从而反过来对孩子未来的收入产生巨大影响。毫无疑问，父母的一些行为是有遗传基础的，比如物质成瘾行为、抑郁症等疾病都有基因因素在起作用。但是有关双胞胎的研究显示，从整体来看，基因因素在预测未来收入方面只起到了一小部分作用。遗传行为的其他来源是父母传递给孩子的文化准则和价值观，因此结论就是传承的文化会对个人经济收益产生很大的影响。[56]

这个结论并不足为奇。如果文化行为的微观准则在组织机构和社会的经济效益中发挥着至关重要的作用，那么人们期望它们在个人经济收益方面发挥作用也是合理的。文化主要传承自父母、对应时期的社会网络以及当地社区这一观察结果提供了一个明确的机制，至少在一定程度上解释了缺乏社会流动性的原因。如果高收入阶层的准则是鼓励个体表现、合作和创新，那么这些准则就会被传承给子女。子女就会运用这些准则去获取经济上的成功，然后再把准则传承给他们的子女。同样，低收入阶层的人可能会陷入文化准则的泥潭，几代人会一直保持着反社会、反经济的行为。20世纪60年代，相当知名的参议员丹尼尔·帕特里克·莫伊尼汉（Daniel Patrick Moynihan）把“贫困文化”认定为美国经济中持续存在下层阶级的主要原因。

我们还可以对这个逻辑进行拓展，用来论证反社会文化准则会导致经济网络中的社会资本变得薄弱。[57]因此，来自文化失调的环境中的个体也较少受益于他们的朋友、邻居和同事。[58]他们很少有机会参加合作

活动，很少进行知识交流，也很少有机会共同承担风险。愿意为年轻人投资的榜样和导师也很少。在这样的文化中，仅有的那点社会资本可能存在于相对消极的团体中，比如帮派之间和毒贩交易网络中。相反，那些来自亲社会文化的人会有完全相反的体验。他们会极大地受益于拥有高社会资本的经济网络的帮助。在这样的网络中，榜样会对年轻人进行投资，人们可以利用自己的关系网为彼此寻找工作，成功的时候大家一起庆祝，遇到困难大家一起分担。

罗尔斯的逻辑与政策

面对左、右派这样的分歧，我们该何去何从？右派认为个人责任仍然很重要，而且积极行为通常会带来积极的结果，这样的观点是正确的。文化规范也许可以解释某些行为的遗传性和持久性，但它们在道德层面无法为反社会活动找借口。同样，尽管美国的社会流动性可能低于一般水平，但它仍然是一个具有流动性的社会，尤其是对中产阶级而言。

左派也有一些正确的观点。与生俱来的财富、智慧加上种族确实在解释经济成果方面发挥着重要作用（即使每一个单项的贡献都非常微小）。如果加入文化因素，出身好坏这一论点就会变得更有说服力，毕竟一个人无法选择自己出生在什么样的文化当中。

从政策角度看，左派和右派常用的解决方案都不理想。收入再分配并不能解决任何行为问题，不管这些行为源于基因还是文化，而且如果对市场采取完全放任自由的管理方式注定会让许多人一辈子贫穷，所以这种方式似乎并不合适。

米勒认为，我们应该放弃左、右派看待这一问题的思想框架，转而采取哲学家约翰·罗尔斯（John Rawls）的立场。[59]我们应该问自己这样一个问题：如果不知道自己的出生是随机抽签决定的，我们想要一个什么样的制度？这样的话，不管我们将出生在上东区一个富有的投资银行世家，还是出生在布朗克斯区一个有毒瘾的单身母亲的家庭，我们都会在知道出生地之前就设计好这个制度。

米勒认为，这个思想实验的答案是，我们会设计一个系统，将顶层的机会均等与底层的社会保障体系结合在一起。首先，我们希望人们有尽可能多的机会走出布朗克斯区，但同时，如果我们运气好，出生在银行世家，也不希望因此受到惩罚。重点在于帮助穷人致富，而不是从经

济上惩罚富人。特别是，我们要确保布朗克斯区的学校和银行家的孩子就读的学校一样好。我们还希望确保孩子的人生道路上不存在不公平的障碍，比如种族歧视。与此同时，我们希望能有对底层人的保护。尽管穷人的孩子已经在学校努力上学，但他或她还需要一些给妈妈做治疗用的康复药物，需要有人帮忙做饭、解决住房及医疗问题，以避免陷入绝境。同样，如果出生在银行世家的孩子知道即便家人运气不好导致倾家荡产，他们的生活也能有基本保障，那么他或她也许可以睡得更安稳些。

米勒把他的罗尔斯逻辑转换成了四点具体的政策建议。第一是提供税收补贴，创建全面的医疗保健体系，使大家能够从提供私人健康计划的保险公司购买医疗保险。第二是大幅度提升教师薪资，实现公共教育质量的提升，尤其是提升针对贫困儿童的教育质量，以换取工会同意实行绩效工资的措施。第三是通过自由党和保守党之间的“大交易”来推进教育改革，允许实行教育凭证制度式的竞争使教育支出与财产税脱钩，这就确保了富裕家庭的孩子能够享受最好的学校教育，并且提高对教育的总体投资。第四是联邦要确保大家的最低生活工资，设置最低收入标准。米勒建议可以挪用美国国内生产总值的2%用于推进这些新举措。

米勒的建议也有利于唤起我们的强互惠公平意识。这些举措的实施定会给品行端正、努力工作的人带来回报，懒惰的人除外。那些运气差的人也能因这些建议举措被慷慨相待。政治家们注意到，这种强互惠的呼声使得米勒的建议受到了大众的广泛支持。

改变文化，创造一个“共同层”

尽管米勒的建议有助于排除前进的障碍，保护社会中某些最脆弱的群体，却无法解决导致贫困和不平等现象出现的更深层次的文化基础。文化与贫困的问题是一个富有挑战性的问题，尤其是当它涉及敏感的种族问题时。20世纪60年代，被多数人视为自由主义者的民主党参议员莫伊尼汉由于讨论贫困文化被左派看成是种族主义者，遭到了猛烈抨击。事实上，讨论贫困文化并不意味着对特定的少数群体的文化进行批判，也并不意味着要让这些少数群体“像白人一样做事”。就像日本人和挪威人所遵从的勤奋工作、合作和创新的准则在本质上没有区别一样，这些文化属性本质上也没有黑种人、白种人、黄种人之分。非洲裔美国籍的总裁、高级政府官员、知名学者及专业人士队伍的不断壮大恰恰证明了

这些准则和行为的效力。这些人的子女往往也非常成功，这又进一步证明了这一点。[60]

种族和文化之间的联系一直在变动，而文化和历史之间的联系是存在的。我们在帕特南对意大利北部与南部所做的对比研究中看到了这一点，甚至在只有合作主体与非合作主体组成的简化世界中也看到了这一点。可以这么说，尽管种族主义对美国社会流动性的阻碍已经减少，但长达250年的奴隶制的遗留问题，再加上随后100年的种族孤立，还是给许多非洲裔美国人的社区文化留下了烙印。现实问题是，那段历史在文化中创造了怎样的反社会和反经济准则，我们又该如何改变这种准则？这种观点调和了非洲裔美国政治领袖和宗教领袖中两大对立阵营的关系，即"指责社会"的左派人士和主张"个人责任"的右派人士。大家可以在承认过去犯过错误的同时，期待准则和个人行为所发生的变化能够成为摆脱贫困文化的关键。

这种观点所传递的积极信息是，文化能够而且确实发生了变化，甚至在几代人身上都发生了变化。而且，文化可以以鼓励经济发展又不牺牲其他积极准则的方式进行变革，使其具有独特性。比如，自20世纪70年代以来，西班牙和爱尔兰都经历了巨大的文化变革，经济增长更是破了纪录。但这两个国家都保持了原有的西班牙传统和爱尔兰传统。消除贫困文化需要变革文化准则，但并不要求美国多样的文化实现完全统一。

话虽如此，但是如果有广泛的共同准则，所有社会都会运作得更好。在文化统一的社会中，人们之间的信任度往往更高，也更容易建立合作。美国和拥有庞大移民社区的欧洲国家所面临的一个日益严峻的挑战是如何在多种族、多元化的文化社会中建立信任和合作。在理想世界中，大家会对成为美国人、英国人、巴西人究竟意味着什么达成共识，并被社会广泛接受。这种共识包括支持民主（例如言论自由权、参与政治进程的重要性）和经济成就（努力工作和创新的回报，受教育和自我提高的重要性）的准则。正是同一文化中所达成的共识，让整个民主国家紧密团结，与贫困文化背道而驰。而共识背后是让多元文化社会展现活力的各种准则、传统和信仰。

创建共识是一项跨越政治体系、教育体系和媒体的挑战。但随着太多公民赶不上经济增长、移民人数增加和全球气温上升导致的社会关系日益紧张，这一挑战变得越发重要。在许多方面，美国都是在创造这种共识方面做得最成功的国家。除了历史上美国对美洲土著人和非洲裔美

国人的排斥之外，它在让移民融入当地社会并带来经济增长和机会方面无可匹敌。美国的开国元老们所创建的不仅仅是法律文件，更是体现价值观的文件。这些价值观为美国服务了两个世纪，但在面临社会剧变的某一时期，在美国国内的家庭、学校、宗教场所、媒体和政治机构中传播、更新和强化这些价值观的机制却败落了。

当今世界，振兴和更新文化可能是美国在新的世纪面临的最重要的任务之一。但这又不仅仅是美国面临的挑战，随着23万信仰伊斯兰教的民众的融入，随着欧盟到底是文化共同体还是政治经济共同体这一答案的确定，欧洲也面临着类似的问题。同时，广大发展中国家的高速发展无疑会带来一场痛苦的文化变革。正如前面所说，倘若没有文化变革，即便得到了援助、有了投资、进行了结构改革，世界上最贫穷的国家还是不会出现经济增长。

同样，这些问题都不可能刚好就是左、右派之间存在的分歧。正如莫伊尼汉参议员总结的那样：“保守主义的核心真理是，决定一个社会成功与否的是文化，而不是政治。自由主义的核心真理是政治可以改变文化，并使之获得拯救。”[61]

未来方向

本章中，我对复杂经济学在公共政策等重要问题方面的应用进行了抽样研究。不管怎样，复杂经济学都可能会帮助我们加深对其他几个领域的理解。

复杂经济学也许无法帮助我们更好地预测通货膨胀，但正如前面所说，它会让我们对商业周期的动态性有新的认识，从而更有效地管理宏观经济政策。一个长期存在的谜题是，为什么工资水平具有黏性，为什么我们无法通过降低工资来阻止经济衰退？在一个完全理性的社会中，我们可以通过小幅削减员工工资，而不是大规模的裁员来消化经济的衰退。强互惠行为也许可以解释这一难题，因为员工认为降薪是不公平的，而仅凭员工那点可怜的动态直觉也无法意识到经济衰退会导致失业率上升。[62]

环境问题是另一个相关的例子。自然环境本身就是一个复杂适应系统。大约开始于1750年的全球财富激增导致大气中的碳排放量达到了高峰。从长期来看，全球财富快速增长所产生的影响已经让全球生态系统产生了戏剧性的、脉冲式的改变，而这些都是由人为因素引起的。我们

正在盲目地对环境进行试验，试验结果目前还无法得知。由于复杂系统往往具有临界点，而且会经历突然的变化或者崩溃，我们应该高度重视这个试验。据某次大规模调查发现，人类对地球的破坏已经超出了它的再生能力，人类所消耗的资源已经是地球承载能力的1.2倍。[63]复杂经济学以自己的方式为此贡献了一分力量，它帮助我们更好地了解了经济和环境的共同进化，重新建立了经济学和热力学的联系，让我们深入了解了为什么人类对这一全球性问题产生的反馈会如此之慢。[64]

各领域对复杂经济学的研究仍在进行，包括医疗体系改革、竞选筹款改革、国际贸易、取消工业管制等领域的研究。[65]这些领域的共同之处是，它们都需要理解复杂适应系统正在经历的变化。

复杂经济学在公共政策领域还是一个非常新的参与者，我在本章中讨论的许多想法也仍处于推测阶段，但新的研究每天都在进行，复杂经济学的研究进展以及大家对它的理解也都在加速当中。

玛格丽特·撒切尔夫人曾说过，“世界上没有‘社会’这种东西，有的只是作为个体的男人和女人，还有家庭”。[66]从复杂经济学的角度看，她完全错了。数百万人进行的互动，包括做决定、参与强互惠行为、执行文化准则、合作、竞争、过日常生活，创造了一个我们称之为社会的新现象，这个现象就像突然出现的漩涡一样真实。社会内部有国家、市场和社区三个概念，这三个要素共同创造了人类生活的经济世界。也许我们无法控制或预测社会的未来走向，但可以努力确保这三者共同发挥作用来创造财富、社会资本和机会。

我认为，复杂经济学超越甚至摒弃了传统的左、右派分类。它不是两极之间模糊的中间地带，而是一个全新的视角，一个看问题的新维度。我个人的希望是，将经济学建立在更加科学的基础之上，让复杂经济学帮助减少美国政治辩论中某些激烈的党派之争，尤其是自20世纪90年代以来的党派之争。毕竟，公民们感兴趣的并不是为其服务的政策制定者中谁的得分更高或者左、右派之间谁胜出，他们只希望看到自己的生活得到改善。

有两种制度为经济发展提供了基础：市场和科学。在此，我们还应该增加第三种：民主。民主本身就是一种政策思想的进化体系。正如爱德华·摩根·福斯特（E. M. Forster）所呼吁的，“请为民主喝彩，这不是因为它承认多样性，而是因为它认可批判主义”。[67]未来几年，复杂经济学将使政治和政策相关的辩论变得更加多元化，最终将由发展而来的进化机制来选择和传播那些最能服务社会的思想。

后记

马赛族流传着一个传说，在18世纪中期，一位族长梦见一群白色的鸟毁灭了马赛族族人的土地。这位族长还梦到巨蛇从海中爬到陆地上吞食自己的族人。马赛族人认为白鸟预言了他们与欧洲殖民者的遭遇，而巨蛇则预言了蒙内铁路的修建。自从这些力量来到家门口，在长达一个世纪的时间里，马赛族人都在不遗余力地保护本族古老的游牧生活方式。

坐在马赛族老人的小屋中时，我在不停地思考经济进化的进程是如何将我们带向不同地方的。一方面，他的生活充满了艰辛。当一位村民给我们看他被感染的伤口时，我更是深刻体会到了这一点。而我可以用抗生素防治感染，任何现代人花几美元就可以做到，而对于那位村民来说，由于没有抗生素，他可能会失去自己的手指，甚至会死去。然而，虽然生活这么艰辛，但我发现当地村民们的脸上洋溢着幸福的笑容，他们的居住环境美丽古朴，邻里关系和睦友爱。相反，当我回到伦敦，我只能注意到地球上最富有的一些人满脸愁容。

我很好奇，经济进化将来会把我们带向何方。我很好奇，这位马赛族长老会做什么样的关于人行道、手机和电视的梦。马赛族人能抵制这些变化多久呢？当变化不可避免地出现，马赛族人会适应吗？或者，他们会丧失本族的立身之本吗？最重要的是，这些变化会使他们更幸福吗？

这些问题不仅关乎马赛族人，更关乎我们所有人。经济进化和增长无疑带来了益处，富有的西方人很容易忘记八代人之前的生活是什么样的，而现在世界上超过1/3的人口还在过着那样古老的生活。过去两个世纪的经济爆炸式增长将数十亿人从饥寒交迫、疾病缠身的霍布斯式存在中解放出来，[1]接下来一个世纪还能再解放数十亿人。但是增长也是有代价的，比如古老的生活方式的消失和环境遭到的破坏。

我们更趋向于认为进化等同于进步，不论是从单细胞生物到人类的物种跨越，还是从石器时代到现代社会的经济跨越。但是进化理论家提醒我们，进化无法完全等同于进步。进步这个术语本身就十分主观，我们只能客观地说，在特定的条件下，进化能随着时间的流逝通向更高的复杂度，而在经济语境中，这意味着更多的财富。但是，理论家还告诉我们，通向更高复杂度的趋势也是不确定的：生物进化历史充满了失败和灭绝，人类社会的历史亦是如此。[2]

我们无法保证经济可以一直停留在1750年以来搭乘的爆炸性增长曲线上。我们必须记住，这样的增长是近期的现象，在人类经济史中只占0.01%。想要增长持续下去，我们就需要运用和发展社会技术，因为它们奠定了增长的基础：市场、科学和民主。即使是在社会技术深度植入的富有的西方，它们也不应当被认为是理所当然的。例如，美国在前段时间陷入了商业丑闻中，并引发了民众对市场的抵制、科学教育的下滑以及民主竞争力的丧失（在2004年的总统选举中，98%二次参选的国会成员都入选了）。[3]经济进化依赖于这些社会技术的健全，它们的持续更新应该成为首要任务。

持续增长曲线还需要市场、科学和民主在世界其他地方植根更深。目前的两个经济飞速增长的大国——印度和中国，都在这方面取得了进步，但是还需要做得更多。印度的民主力量十分强大，科学也进步不小，但是市场改革不尽如人意。与此同时，中国在科学方面也作出了较多的投资。令人感到悲哀的是，在世界大部分地区，尤其是撒哈拉以南的非洲，基本上还未被这三项社会技术影响。在这些地方，腐败、迷信和强权依然大行其道。有条不紊地建立经济进化所需要的制度和文化是这些国家摆脱贫穷的唯一途径。

不过，我们也有理由对增长曲线的持续持乐观的态度。此前，我介绍了马文·明斯基，并且讨论了组成“心智社会”的人擅长的非凡技能。在全球范围内，经济都在进化为一个心智社会。物理技术的革新持续以指数速度降低人与人之间的互动成本，提高互动效率。[4]几乎整个人类知识库都在被数字化，向地球上任何地方的任何人敞开大门。从根本上说，全球心智社会正发展出全球记忆。物理技术的革新以及全球心智的增强，将极大地增加系统的可能性。我们无法想象这样的系统有多么强大。

持续的进化以及物质技术和社会技术的扩散给了我们乐观的理由。一些人可能会争辩说，对人类来说，真正的风险不是从增长曲线上掉落，而是继续留在上面。这里存在三个问题。首先，像弗兰肯斯坦博士的怪物一样，全球经济可能已经开始攻击它的创造者了，它吞噬着地球的资源，污染了地球的土地、海洋和天空。如前所述，环境本身是一个复杂适应系统，临界点、巨变，甚至是崩溃的风险都是真实存在的。其次，自18世纪以来，物理技术进化的节奏已经远远超过了社会技术。我们已经在与核技术和基因技术的影响博弈，在下一代人身上，人工智能与纳米技术也可能会被列入博弈的名单。如果我们的社会技术无法迎头

赶上，全球灾难来临的风险将会持续增加。最后是文化的碰撞。[5]此前，文化冲突只存在于地理边界。如今，文化冲突每天都在电视、网络以及多元文化城市中发生。不幸的是，人类在控制文化冲突方面——从科尔特斯人入侵墨西哥事件到伊斯兰与西方世界的冲突，并不令人满意。[6]

虽然存在这些挑战，但在我看来，天平还是倾向于乐观，因为我们现在有能力理解复杂的人类系统。我们可能无法预测或者控制这些系统，但是可以通过自己的行动来塑造它们。

毫无疑问，不论是政治领袖还是商业领袖都需要愿景和智慧来采取许多行动以解决这些问题。但是我们应该记住，在进化系统中，力量并非自上而下，而是自下而上的。进化是一个盲目的过程，而进化算法会回应接收到的任何适合度函数。如果作为个体消费者、工人和选民，我们要求所处的经济体制和政治体制最大化满足自己的短期需要，用更多的东西来填充物质生活，不考虑地球和未来子孙的健康，那么我们将得偿所愿。

但我们并非别无选择。通过选择消费方式、雇主、候选人以及自己的声音，我们能创造出另一种适合度函数，它需要我们的商业系统、政府和科研机构采取长远的视角，用更广阔和可持续发展的方式来解决全球社会的需求。如果我们创造出这种适合度函数，那么那些机构和经济就必然会适应和响应这一号召。埃德蒙·伯克（Edmund Burke）曾说过，社会是一种“关系，它不仅仅是所有生者之间的关系，还是所有生者、死者和未出生的人们之间的关系”。[7]我们都是全球心智社会的一部分，社会将怎样进化取决于我们。

注释

序言

1．参见Kuhn1962。引自尼古拉斯·韦德对库恩的科学研究的回顾，摘自*Science*, vol. 197, July 8, 1977, pp. 143-145。

2．人们普遍认为是托马斯·卡莱尔使用“沉闷的科学”一词来形容马尔萨斯的理论：Samuelson and Nordhaus 1998, p.323, or Heilbroner 1953, p. 76。这实际上是不正确的。卡莱尔首先使用这个短语来进一步表述他自己的令人震惊的种族主义观念。卡莱尔在1849年出版的一本小册子上发表了这番话，这本小册子被整理成为“*An Occasional Discourse on the Negro Question*”。在这期间，经济学家们，例如斯图亚特·密尔正在对奴隶制进行声讨，称其否定了基本的人权。当卡莱尔使用这些词汇时，他是在嘲笑经济学家和废奴主义者的联盟。

3．John Maynard Keynes’s 1936 *The General Theory of Employment, Interest and Money*, ch 24, part V.

4．Brown and Eisenhardt 1998, Clippinge 1999, Kelly and Allison 1999, Lis-sack and Roos 1999, Petzinger 1999, Koch 2000,Lewin and Régine 2000, Pascale, Millemann, and Gioja 2000,and Stacey Griffin, and Shaw 2000.

01

1．Smith 1776, chapter IV, p. 25.

2．Kahneman, Diener and Schwarz 1999.

3．Krugman 1992.

4．Kauffman 1995. Kauffman 2000, pp. 211-241.

5．Anderson, Arrow, and Pines 1988.

6．Gell-Mann 1994 and Flake 1998. Haken 2000.

7．Seabright 2004, pp. 13-26.

8．Diamond 1997, Wright 2000, Landes 1998, Jay 2000, Cameron and Neal 2003, and Seabright 2004.

9．ones, et. al. 1992, and Diamond 1997.

10．Horan, Bulte, and Shogren 2005.

11．Seabright 2004. Bonabeau, Dorigo, and Theraulaz, eds. 1999.

12．Horan, Bulte, and Shogren 2005.

13．Chagnon 1992.

14．www.j-bradford-delong.net, accessed on July 16, 2002. New York state government statistics.

15．Chagnon 1992.

16．www.j-bradford-delong.net, accessed July 16, 2002.

17．Schwartz 2004, pp. 9-22.

18．Chagnon 1992.

19．Petroski 1992,23. Schwartz 2004 pp. 9-22.

20．Diamond, 1997, pp. 45-50.

21．德龙经由七组数据构建了这一结果：克莱默对人口规模的一组长期估计，三组人均GDP的长期数据，以及世界GDP的三组数据。这个几组数据和关于德龙如何构建出这一结果的描述发表在德龙的网站上，名为“公元前100万年前至今的世界GDP估算”。

22．Bernstein 2004, pp. 17-23.

23．Ormerod 1994, p. 10.

24．Landes 1969, p.5.

25．Kauffman 2000, p. 211.

26．Dennett, 1995, pp. 48-50. Dawkins 1976 and 1982 and Dennett 1995.

27．Rothschild, 1990, “The Power of Biobabble” published on the *Slate* website October 24,1997.

28．Holland 1992, Whitley, ed. 1993, Mitchell 1996, Landweber and Winfree, eds. 2002, and Crutchfield and Schuster, eds. 2003.

29．Hodgson 1993, p. 81.

30 . Dennett 1995, pp. 28-34; 48-60.

31 . Dembski and Ruse, eds. 2004.

32 . Kauffman 1993. Wright 2000, and Cameron and Neal 2003.

33 . Dennett 1995. Kauffman 1995, pp. 149-189.

34 . Seabright, 2004.

35 . Nelson 2003.

36 . 社会技术与经济学家所说的制度近似，但又不完全相同，我将在第12章更加全面地定义社会技术。

37 . Freeman and Soete 1997.

38 . Hodgson 1993.

39 . Charles Darwin's Autobiography, p. 120, as quoted in Plotkin 1993, pp. 28-29.

40 . Browne, 2002.

41 . Thorstein Veblen, "Why is Economics Not an Evolutionary Science?", *The Quarterly Journal of Economics*, XII 1898, pp. 373-397, reprinted in Gherity 1965.

42 . Hodgson 1993, Chapter 7.

43 . Hodgson 1993, Chapter 12. Vriend 2002. Colander, ed. 2000.

44 . Nelson and Winter 1982.

45 . Waldrop 1992 pp. 82.

46 . Waldrop 1992. Gell-Mann 1994, Kauffman 1995, Holland 1998, and Johnson 2001. Flake 1998. Bar-Yam 1997. Cowan, Pines and Meltzer, eds. 1994.

47 . Anderson, Arrow, and Pines, eds. 1988, and Cowan, Pines and Meltzer, eds. 1994.

48 . Arthur 1999.

49 . Mark Blaug's essay in Hausman, ed. 1994, pp. 348-375. Hands 2001.

02

1. Waldrop 1992.

2. Waldrop 1992, p. 91.

3. Keen 2001, Fullbrook, ed. 2004, and Ormerod 1994.

4. Fullbrook, ed. 2004.

5. Colander, Holt, and Rosser, eds. 2004.

6. Cassidy 1996.

7. Alan Greenspan quoted in Andrews 2005.

8. Hausman 1994 and Hands 2001.

9. Niehans 1990 and Backhouse 2002. Samuelson and Nordhaus 1998 and Stiglitz 1997.Mansfield 1999.

10. 一些读者会注意到，我对传统经济学共识的描述具有明显的英美新古典主义的偏差。这是因为近几十年来这些想法在全球占主导地位。大陆学者通常比英美学者更重视历史和制度经济学的观点。稍后我将讨论这些观点与复杂经济学的相关性。

11. Nelson and Winter 1982, pp. 6-11.

12. Samuelson and Nordhaus 1998, Stiglitz 1997, and Mas-Colell, Whinston, and Green 1995.For macroeconomics: Dornbusch and Fischer 1990, Mankiw 1994, Krugman and Obstfeld 1991, Romer 1996, Blanchard and Fischer 1989, and Heijdra and Van Der Ploeg 2002.Aghion, Frydman, Stiglitz, and Woodford, eds. 2003 and Szenberg and Ramrattan, eds.2004.

13. Baumol 2000, pp. 3-4.

14. Szenberg and Ramrattan, eds. 2004.

15. 并非所有的诺贝尔奖得主都属于我所谓的传统阵营，例如赫伯特·西蒙、弗里德里希·哈耶克、道格拉斯·诺斯和丹尼尔·卡尼曼的工作都为复杂经济学家的工作提供了基础。

16. Colander, ed. 1999, pp. 6-7.

17. Backhouse 2002, pp. 13-17.

18 . Adam Smith are from Ross 1995, Niehans 1990, and Backhouse 2002. Niehans, 1990, pp.9-13.

19 . Niehans, 1990.

20 . Smith 1776, p. ix.

21 . Backhouse 2002, p. 124.

22 . Smith 1776, Book I, Chapter 1, p. 4.

23 . Niehans 1990, p. 60.

24 . Smith 1776, Book I, Chp. 2, p. 15.

25 . Smith 1776, Book IV, Chp. 2, p. 482.

26 . mith 1776, Book IV, Chp. 2, p. 485.

27 . Niehans 1990 and Backhouse 2002.

28 . Niehans 1990, pp. 24-36.

29 . Mirowski 1989.

30 . Louis XIV. Backhouse 2002, p. 91.

31 . Backhouse 2002, pp. 104-108, Niehans 1990, pp. 73-76.

32 . 我们将在第二部分继续讨论。

33 . Antoine Augustin Cournot and Johann Heinrich von Thü nen. Niehans 1990, pp. 164-187.

34 . Niehans 1990, pp. 123-126.

35 . Niehans 1990, pp. 118-137, and Backhouse 2002, pp. 132-165.

36 . Backhouse 2002, p. 136.

37 . Niehans 1990, pp. 187-196.

38 . Ingrao and Israel 1990, p. 87.

39 . Igrao and Israel 1990, p. 88.

40 . Mirowski 1989 Niehans 1990, and Backhouse 2002.

41 . Backhouse 2002, pp. 237-240. Backhouse 2002, pp. 166-168.

42 . Stewart 1989.

43 . Ingrao and Israel 1990.

44 . Stewart 1989, p. 60.

45 . Ingrao and Israel 1990, p. 88 and Mirowski 1989, pp. 219-220.

46 . Ingrao and Israel 1990, p. 88.

47 . Mas-Colell, Whinston, and Green 1995, pp. 584-598.

48 . Mirowski 1989, pp. 243-248.

49 . Niehans 1990, pp. 197-207.

50 . Niehans 1990, p. 197 and p. 198.

51 . Mirowski 1989, p. 217, pp. 256-257.

52 . Mirowski 1989, p. 257.

53 . Mirowski 1989, p. 219.

54 . Mirowski 1989, pp. 217-222. Niehans 1990, p. 189, p. 201.

55 . Mirowski 1989, p. 219.

56 . Niehans 1990, pp. 259-266.

57 . Niehans 1990, p. 265.

58 . Voltaire's *Candide*, 1759, Chapter I.

59 . Mirowski 1989, pp. 219-221.

60 . Niehans 1990, pp. 420-444.

61 . Backhouse 2002, p. 259.

62 . Mas-Colell, Whinston and Green 1995, pp. 5-16.

63 . Mas-Colell, Whinston and Green 1995, pp.691-693.

64 . Hayek 1988.

65 . Niehans 1990，第445-451页，约翰·埃利奥特在1983年版Schumpeter 1934中的介绍。

66 . John Elliot in Schumpeter 1934, 1983 edition, p. xix.

67 . John Elliot in Schumpeter 1934, 1983 edition, p. xxiv.

68 . Nelson 1996, Scherer 1999, and Helpman 2004.

69．Niehans 1990, pp. 451-456.

70．罗伯特·索洛偶尔也会担任麦肯锡公司的顾问一职。

71．Solow 2000, pp. ix-xxvi.

72．Solow 1956.

73．Solow 2000, Barro and Sala-i-Martin 1995 and Romer 1996.

74．Romer 1994, Aghion and Howitt 1998 and Barro and Sala-i-Martin 1995.

75．Romer 1990.

76．Heijdra and Van Der Ploeg 2002.

77．应用最广泛的迈克尔·波特关于战略的观点源于新古典主义微观经济学，参见Porter 1980 and 1985。

78． Niehans 1990, p. 491.

03

1． Waldrop 1992.

2． Anderson 1972.

3． Haynie 2001.

4．我所属的麦肯锡公司自1994年以来一直是圣塔菲研究所的研究项目的资金支持者。

5． Waldrop 1992, pp. 136-197, Anderson, Arrow and Pines, Eds. 1988.

6． Waldrop 1992, p. 142.

7． Ingrao and Israel, p. 158.

8．Ingrao and Israel, p.158.

9． Mirowski 1988 pp. 241-265 and Ingrao and Israel 1990, pp. 148-173.

10．Friedman 1953.

11．Hands 2001, p. 53. Hausman, 1994.

12. Simon, H. 1963. "Testability and Approximation", *American Economic Review: Papers and Proceedings*, vol. 53, reprinted in Hausman, ed. 1994, pp. 214-216.

13. 这是我简化后的结果，但逻辑与西蒙的例子是一样的。西蒙指出，纳格尔证明了使用Z的有效性来支持X和Y的谬论，并且进一步认为，即使观察到Z，也需要观察X和Y来支持这一理论。

14. "Why Look Under the Hood?" Hausman, ed. 1994, pp. 217-221.

15. Hands 2001, pp. 53-60. Hands 2001. Hausman 1994.

16. Hands, 2001, p. 311. Holland 1998, pp. 28-33. Sterman 2002.

17. Axel Leijonhufvud, "Towards a Not-Too-Rational Macroeconomics," Chapter 3, pp. 39-55,in Colander, ed. 1996.

18. Kahneman, Slovic, and Tversky 1982.

19. 技术在确定满足行为的界限方面起着重要作用。例如，我们不可能在未来让附近所有加油站通过无线互联网将油价发送到汽车上，然后通过GPS定位最便宜的加油站，车载电脑还要计算额外的行驶距离和成本。即使有更多的可用信息，人们还是可能会关闭计算机，不去费心计算可以节省几美元，而较近的车站可能有更干净的浴室。

20. Camerer 2003, pp. 473-476.

21. Kirman 1999 p. 10. Hicks, 1939, p. 122.

22. 假设有一块区域在随机地抛撒20美元的钞票。该地区的人口分布不均衡，人们主要聚集在一个我们称之为A的地区。在某一点上，钞票会在一个相对不受欢迎的地区（称之为B）积累起来。一旦发现这种现象，人们就会冲到B区去抢钞票。但由于这需要时间，当人们离开的时候，钞票会在A区积累起来，于是，人们又会冲回A区，这种变化造成了整体格局上的波动。

23. Richard Day 1994 and 1999. Fudenberg and Tirole 1991 and Heijdra and Van Der Ploeg 2002.

24. Anderson, Arrow, and Pines, eds. 1988, Arthur, Durlauf, and Lane, eds. 1997, Sterman 2000,and Durlauf and Young, eds. 2001.

25 . Aghion and Howitt, 1998.

26 . Crutchfield and Schuster, eds. 2003, pp. 65-78, and pp. 81-100.

27 . Cutler, Poterba, and Summers 1989.

28 . Mankiw 1994, p. 326.

29 . Mordecai Kurz 1997.

30 . Arthur 1994a.

31 . Keynes 1923, p. 65.

32 . Thomas Love Peacock's 1835, *Crotchet Castle*, Chapter II: The March of Mind.

33 . Hempel 1965. Hands 2001, pp. 82-88 and Hausman, ed. 1994, pp. 1-50.

34 . Klemke, Hollinger, and Kline, eds. 1980, pp. 19-34. Hands 2001, pp. 88-93.

35 . Hull 1988.

36 . Kirman and Gé rard-Varet 1999, p. 8.

37 . Greene 2000.

38 . Campbell, Lo, and MacKinlay 1997, Mandelbrot 1997, Lo and MacKinlay 1999, Mantegna and Stanley 2000, Shleifer 2000, Shiller 2000, Johnson, Jeffries, and Hui 2003, and Sornette 2003.

39 . Kagel and Roth, eds. 1995, Camerer, Loewenstein, and Rabin, eds. 2004.

40 . Blanchard and Fischer 1989, pp. 301-308, 332-336. Sterman 2000, pp. 661-842.

41 . Lillo and Farmer 2004.

42 . Samuelson and Nordhaus 1998, p. 681.

43 . Cave 2004.

44 . Montier 2002, pp. 29-31.

45 . Nagel, Shubik, Paczuski, and Bak 2000.

46．“今天的宏观经济学牢固地建立在一般均衡结构中，现代模型将经济描述为处于暂时的均衡状态。”参见Blanchard 2000, p. 1402。

47．Scarf and Hansen 1973 and Axtell 2002, 2003.

48．Scarf and Hansen 1973.

49．Lo and MacKinlay 1999, pp. 26-40.

50．Mandelbrot 1997, Lo and MacKinlay 1999, Mantegna and Stanley 2000, and Johnson, Jeffries, and Hui 2003.

51．Sornette 2003 and Mandelbrot and Hudson 2004.

52．更完整的讨论见第17章。

53．Ijiri and Simon 1977，感谢多因·法默让我意识到这句话。

54．Greene 1999.

55．关于隐喻和科学模型之间关系的讨论，参见Holland 1998, Cartwright 1999 and Hands 2001。

56．Mirowski 1989 and Ingrao and Israel 1990. Mirowski 1989, p. 224.

57．Anderson, Arrow, and Pines, eds. 1988.

58．Baeyer 1998, Kondepudi and Prigogine 1998.

59．Mirowski 1989, pp. 230-233, p. 314, 328. Smith and Foley 2002.

60．如前所述，直到20世纪50年代，新古典主义经济学和增长理论才得以统一。然而，正如我们将在第三部分看到的，新古典主义增长模型仍然有一个如何实现真正的创新的问题。

61．Lionel Robbins’ *An Essay on the Nature and Significance of Economic Science*, reprinted in Hausman 1994, pp. 83-110.

62．Samuelson and Nordhaus 1998, p. 4.

63．Bayer, 1998, p. 133.

64．Baeyer 1998, p. 56.

65．我找不到1867年出版的汤姆森和泰特的书的原版。大英图书馆中有1890年出版的第四版复印件，汤姆森和泰特在书中详细介绍了不同版本之间的差异。可以确定的是，1890年版的书中仍然没有提到

熵。

66．多宇宙物理学理论有时会提出关于宇宙之间通过黑洞或量子效应相互作用的可能性的问题。

67． Waldrop 1992, p. 147.

68． Mas-Colell, Whinston, and Green 1995, p. 620. Blanchard 2000, p. 1402.

69． Georgescu-Roegen, 1971.

70．当一个封闭的系统达到最小能量平衡时，能量可能发生流动，并在此过程中产生秩序。我们可以把能量能自由流动并产生秩序的区域，视作一个更大的、封闭系统内的一个子系统。

71． Nicholas Georgescu-Roegen 1971 and Jay Forrester 1961, Meadows, Randers, and Meadows, 2004.

72． 1996年，我在访问纳斯达克计算机中心时听到了这个故事。虽然该中心有备用发电机，但很明显，松鼠还是造成了一系列的连锁反应，导致备用电源接通时间发生延迟。人们倾向于对所有发生在重要机构的意外都严肃对待，从那时起，他们的设备已被改进，用于预防此类意外事件。

73． Bender, T. and Schorske, C., eds. 1997. *American Academic Culture in Transformation*.Princeton, N.J.: Princeton University Press, pp.73-74, quoted in Mirowski 2002, p. 8n3.

74． Maynard Smith, 1982.

75． Devlin 2000. Lakoff and Nunez 2000.

76． Ingrao and Israel 1990, pp. 280-288. Alan Kirman in Kirman and Gé rard-Varet 1999,Chapter 1.

77． Deutsch (1997) , ch. 10.

78．这并不意味着只有直观的数学对象才是真实的。物理学家创造了各种奇异的数学物体，这些物体不是直观的，但具有真正的物理学意义。例如，超弦理论假设宇宙中有11个或更多个维度。但归根结底，这种分类和解释与一些与物理世界的相互作用有关，不管是一个复杂的物理实验，还是一个孩子在数鹅卵石。

79． The *Oxford Dictionary of Physics* 2000, pp. 158-159.

80 . The *Collins Dictionary of Economics* 2000, p. 16. Stiglitz 1997, p. 88.

81 . Daly 1999, pp. 75-88.

82 . Anderson, Arrow and Pines, eds. 1988, Waldrop 1992.

83 . Arthur, Durlauf and Lane, eds. 1997, Durlauf and Young 2001, Bowles, Gintis, and Osborne Groves 2005, Gintis, Bowles, Boyd, and Fehr 2005, Farmer and Geanakopolos 2005, and the Santa Fe Institute website www.santafe.edu.

84 . Krugman 1998.

04

1 . Seth Mydans, "In a Philippine Wasteland, a 'Living Metaphor,'" *The International Herald Tribune* reprinted from the *New York Times*, July 19, 2000, p. 1.

2 . Hicks' Value and Capital 1939, Schumpeter 1934, Hayek 1948, North 1990 and 2005,Rosenberg and Birdzell 1986, Mokyr 1990 and 2002.

3 . Epstein and Axtell 1996, Rauch 2002.

4 . Waldrop 1992, pp. 241-243.

5 . Epstein and Axtell 1996, Chapter 2.

6 . Epstein and Axtell 1996, pp. 71-82.

7 . Aghion, Frydman, Stiglitz, and Woodford, eds. 2003, p. 368.

8 . Epstein and Axtell 1996, pp. 54-68.

9 . Epstein and Axtell 1996, pp. 94-137.

10 . Epstein and Axtell 1996, p.97.

11 . Lo and MacKinlay 1999 and Thaler 1993.

12 . Axtell 2002.

13 . Epstein and Axtell 1996, pp. 36-37, p. 154.

14 . Colander, ed. 2000, Mirowski 2002.

15 . Anderson, Arrow, and Pines, eds. 1988, Arthur 1994, Colander, ed. 1996, Arthur, Durlauf, and Lane, eds. 1997, Axelrod 1997, Albin 1998, Lesourne and Orlé an, eds. 1998, Prietula, Carley, and Gasser, eds. 1998, Young 1998, Gintis 2000, Durlauf and Young, eds. 2001,Rauch and Casella, eds. 2001, Camerer, ed. 2003, Johnson, Jefferies, and Hui 2003,Kollman, Miller, and Page, eds. 2003, Bowles 2004, Farmer and Geanakopolos, eds. 2005.

05

1 . Gleick 1987,Stewart 1989,Devaney 1992, Flake 1998, Kaplan and Glass 1995, Strogaz 1994,Bar Yam 1997.

2 . Devaney's 1992, pp. 9-11.

3 . Forrester 1961 ，Sterman 2000 ，Senge 1990.

4 . Sterman 2000, pp. 649-654.

5 . Skidelsky 1994.

6 . 从技术上讲，如果输入和输出之间有足够长的延迟，负反馈可以驱动振幅增大，但负反馈最常见的影响是抑制。

7 . Stewart 1989, pp. 60-72 and Mirowski 1989, p. 72.

8 . Gleick 1987, Stewart 1989, and Devaney 1992.

9 . Stewart 1989, pp. 81-84.

10 . LeBaron 1994.

11 . Peters 1991 and Chorafas 1994.

12 . Brock, pp. 77-97 in Anderson, Arrow, and Pines, eds. 1988, Le Baron 1994, Brock, pp. 385-423 in Arthur, Durlauf, and Lane, eds. 1997, Mandelbrot 1997, Mantegna and Stanley 2000,Sornette 2003, Johnson, Jeffries, and Hui 2003 for discussions.

13 . Stewart 1989, pp. 60-62.

14 . Chapter 20 of Sterman 2000.

15．Sterman 2000, pp. 791-798.

16．Sterman 2000, p. 798.

17．Sterman 2000, pp. 798-841.

18．Sterman 2000, pp. 824-828.

19．Paich and Sterman 1993.

20．Sterman 1989a and 1989b.

06

1．这是一幅合成图像，用来说明各种传统经济模型中的典型理性假设。

2．Pindyck and Rubinfeld 1989，Romer 1996，Plous 1993，Kirman and Salmon 1995.

3．Simon 1997, and Simon 1992.

4．Selten 1990，Akerlof and Yellen 1985.

5．Selten 1990, Nordhaus 1998, p. 178, Stiglitz 1997 or Mas-Colell et. al. 1995, Pindyck and Rubinfeld 1989, Dornbusch and Fischer 1990, or Romer 1996.

6．Rubinstein 1998 and Sterman 2000.

7．Camerer, Lowenstein, and Rabin, eds. 2004, p. 6. Kahneman, Slovic and Tversky 1982.

8．Hogarth 1990, Thaler 1992, Plous 1993, Thaler 1993, Kagel and Roth 1995, Henrich, et. al.,eds. 2004, Montier 2002, Camerer, Lowenstein, and Rabin, eds. 2004. Camerer 2003, and Gintis, Bowles, Boyd, and Fehr 2005. Russo and Schoemaker 1989, Bazerman 1998, and Dö rner 1996.

9．“Rethinking Thinking,” *The Economist*, December 18, 1999, pp. 63-65, Roger Lowenstein,“Exuberance is Rational,” *The New York Times Sunday Magazine*, February 11, 2001, and Louis Uchitelle, “Some Economists Call Behavior a Key,” *The New York Times*, business section, February 11, 2001.

10．Bazerman 1998, p. 81.

11．Kagel and Roth 1995, pp. 253-348.

12．Fehr and Gä chter 2000.

13．“Primary Sources”, the Atlantic Monthly, April 2003, pp. 33-34.

14．Gintis, Bowles, Boyd, and Fehr, eds. 2005, pp. 3-39.

15．一个传统的经济学家可能会说，互惠行为有一个合理的基础。人们会假设未来他们会有互动，因此这些策略在一个重复进行的游戏中是有意义的。我对此有两点需要澄清：第一，我从来没有说过这种行为是不合理的。事实上，这是一种进化的逻辑。第二，不管将来是否会有互动，人们都倾向于这么做，因为它是启发式行为，而不是计算后的行为。

16． “Rethinking Thinking,” *The Economist*, December 18, 1999, p. 65.

17．Roxburgh 2003.

18．“The Price of Safety,” *The Economist*, November 23, 2000.

19． James Surowiecki, “The Financial Page: Mind Over Money,” *The New Yorker*, April 23 & 30, 2001, p. 60.

20．Fischer Black in Thaler 1993, pp. 3-22.

21．Lewis 1985.

22．Deutsch 1997, pp. 131-140.

23．Steven Pinker 1997, Andy Clark 2001, and Daniel Dennett 1991.

24．Foster and Young 2001.

25．Arthur 1994b.

26．更具体地说，考虑到在行为主体身上施加的信息和协调约束，酒吧问题中的固定点是不稳定和不可访问的，参见Arthur 1994b。

27． Challet and Zhang 1997, Johnson, Jeffries and Hui 2003, pp. 87-136.

28．Stillings, et. al. 1995, Clark 1999, and Clark 2001, Pinker 1997.

29．Pinker 1997.

30. Schank 1990, McKee 2003.

31. Mitchell 1993.

32. Gilovich, Vallone and Tversky 1985.

33. Clark 2001, p. 38.

34. 教电脑从餐馆订餐并不是一件小事，有研究者在20世纪70年代就在尝试这一点，参见Schank 1990。

35. 布莱恩·阿瑟在1992年的圣塔菲研究所工作文件中指出了这种归纳与演绎的对比。

36. Holland, et. al. 1986.

37. Holland 1995, pp. 43-62.

38. Holland, et. al. 1986, pp. 68-76.

39. Schank 1990, pp. 4-5.

40. Mitchell 1993.

41. Arthur 1995, and Arthur, Holland, LeBaron, Palmer, and Tayler, “Asset Pricing Under Endogenous Expectations in an Artificial Stock Market,” in Arthur, Durlauf, and Lane, eds.1997, pp. 15-44.

42. 这条规则与完全理性相似，但并不精确，其原因可以追溯到在图灵机上实现完美理性的问题，人们只能在计算机上创建完美理性的近似值。

07

1. Rauch and Casella, eds. 2001, pp. 3-5, Watts 2003, Giddens 2001.

2. Watts 1999, Chapter 2.

3. Watts 2003 and Barabá si 2002.

4. Kauffman 1995, pp. 55-58, Kauffman 1993, pp. 307-310. Watts 2003, pp. 43-47.

5. Kauffman 1993, pp. 307-310 and 1995, pp. 55-58, Watts 2003, pp. 43-47.

6. Gladwell 2000.

7．Farrell 1998.

8．Watts 2003.

9．Watts 1999, pp. 3-5, and Watts 2003, pp. 92-95.

10．数字510、829是所有有限贝肯系数的男演员和女演员总数。样本中有10%的人与贝肯没有关联，或者有着无穷大的贝肯系数。

11． Watts 2003, pp. 121-129.

12． Watts 2003, pp. 39-40.

13． Watts 1999 and Newman 1999.

14． Newman 1999 pp. 3-4.

15．这个口号归属于升阳公司。

16．它们也是细胞自动机的一种形式，这一概念是由冯·诺依曼提出的。

17． Kauffman 1995, pp. 74-80.

18．布尔网络是一种特殊的元胞自动机。物理学家和数学家斯蒂芬·沃尔夫拉姆等人认为，宇宙本身就是一个由细胞组成的机器人，经济也是这样。所有这些观念都是纯理论性的，关键是利用计算机对经济现象进行分析能够产生有趣的见解，是一个非常庞大的研究领域。

19． IBM “蓝色基因”的运算速度为每秒70.62万亿浮点。

20．然而，如果量子计算机最终被开发出来，无法完全探索大型网络的现状可能会有所改变。参见Johnson 2003。

21． Cookson and Griffith 2001.

22． Kauffman 1993, pp. 191-203.

23．此示例摘自当时的新闻报道、戴尔和IBM的各种介绍文件以及公司的网站。

24． Porter 1996.

25． Kauffman 1993, pp. 209-218.

26． Kauffman 1995, pp. 252-271，Eisenhardt and Brown 1999.

27．在第16章中，我们将研究密歇根大学计算经济学家斯科特·佩奇的工作，他研究了层次结构在解决计算网络难题中的作用，参见Page 1996。

28．KKauffman 1993, pp. 209-227.

29．Kauffman 1995, pp. 84-92.

30．除了对初始条件和扰动的敏感性之外，对于给定数量节点的网络，混沌在布尔网络中表现为更长的中间值状态循环，参见Kauffman 1993, p. 197。

31．Katzenbach and Smith 1993.

08

1．改编自Fischer 1996, pp. 35-37。

2．Kennedy 1987.

3．Moore 1983 and Gordon 1986.

4．Niehans 1990, pp. 372-378.

5．Heijdra and van der Ploeg 2002, pp. 477-539.

6．想象一下，我们把勺子敲击的输入编码成一个字符串，然后通过压缩算法运行字符串并尽可能地压缩它。如果勺子敲击输入真的是随机的，我们将不能大幅压缩。如果我们对输出字符串进行编码，发现它比输入字符串长，这是因为一个输入点可能会产生多个振荡输出摆动。通过压缩算法运行输出时，我们将得到一个长度与原始字符串相等或稍长一点的字符串，这取决于压缩的效率。我们既没有与熵作斗争，也没有向输入字符串添加实际信息，我们只是在被压缩算法过滤掉的噪声中添加了一些“回声”。相反，如果我们选取一个随机的字符串作为输入值，例如01001001011101001，然后运行，即使用能量来对抗熵，字符串将会重新排列成相同长度的输出，并被读取为00000000011111111，这可以进一步压缩为“打印9次0，打印8次1”的声明，这就可以编码成少于17位的东西了。具体参见Gell-Mann 1994。

7．Heijdra and van der Ploeg 2002, p. 528.

8．Skidelsky 1994.

9．克鲁格曼对凯恩斯主义经济学的发展有一个很好的描述，具体参见Krugman 1994。

10．Akerlof and Yellen 1985a and 1985b, Krugman 1994, pp.

206-220.

11．Heijdra and van Der Ploeg 2002, pp. 359-403.

12．Kurz, 1997.

13．Heijdra and van der Ploeg, 2002, p. 402.

14． Holland 1998, Buchanan 2000, Johnson 2001, and Morowitz 2002.

15．Anderson 1972.

16．Nijhout, Nadel, and Stein 1997.

17．俄罗斯经济学家尼古拉·康德拉耶夫认为，俄罗斯经济也有一个长达60年的“长波”周期，具体参见Sterman 1985。

18．Lotka 1956.

19．Sterman 2000, pp. 684-708. Senge 1990.

20．Sterman 1989a, 1989b, and 2000, pp. 684-698.

21．Sterman 1989b.

22． Chapter 9 in Casti and Karlqvist 1991. Sterman 2000, pp. 684-698.

23．在游戏的时间框架内，振荡无法恢复均衡。如果没有某种协调机制，例如允许所有人见面协调订单，振荡也不太可能在更长时间内恢复均衡。即使玩家在听完一场关于游戏动态结构的讲座后，再次使用不同的订单牌进行游戏，仍然无法达到完全的理性均衡。

24．啤酒分销游戏中信息的增加并不会改变完全理性的结果，同样，游戏参与者拥有计算理性均衡所需的所有信息。相反，附加的信息可以识别、锚定和调整启发式带来的问题，并提高启发式的性能。

25．Sensier and van Dijk 2004.

26． McCarthy and Zakrajš ek 2002 for an analysis and further references.

27．有人可能会注意到，我刚刚提到了一个导致商业周期放缓的外生因素。在对传统理论的批判中，我并没有否认外生因素的存在和影响，而是批判了外生因素在理论中的典型作用。在这种情况下，我的论点是，商业周期是内生动态结构的一个涌现属性，因此不是外

生的，但技术变化等外生因素可以影响动态结构本身。

28．Gould and Eldridge 1993 and Gould 2002.

29．Gould 1989, p. 54.

30．Gould 2002, pp. 774-775，pp. 824-839.

31．Bak 1996，Mantegna and Stanley 2000 and Johnson, Jeffries and Hui 2003.

32．关于社会和生物背景下的网络结构的具体讨论参见Girvan, M. and Newman, M.E.J.2001。

33．Jain and Krishna 2002a and Jain and Krishna 2002b.

34．Marx's Critique of Political Economy and Schumpeter's *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process*.

35．Lydia Adetunji, "Inventor found investors not on his wavelength," *The Financial Times*,December 19, 2001, p. 5.

36．Ziman, ed. 2000.

37．Rosenberg 1982, pp. 55-80.

38．Kauffman in Anderson, Arrow, and Pines, eds. 1988, pp. 125-146.

39．Baldwin and Clark 2000.

40．Cutler, Poterba, and Summers 1989.

41．Buchanan 2000, p. 35.

42．这个分布被称为幂律，因为一个变量是另一个变量的指数幂。

43．Bak 1996 and Buchanan 2000.

44．Niehans 1990, pp. 259-266.

45．Mandelbrot 1997, pp. 252-269.

46．Mandelbrot 1997, pp. 371-411.

47．摘自斯坦利于2001年11月16日在圣塔菲研究所举办的"经济作为一个不断发展的复杂系统"第三次会议上的演讲。

48．Plerou, et. al. 1999.

49．Axtell 2001.

50．Stanley et. al. 1996 and Lee et. al. 1998.

51．Farmer, et. al. 2004, Lillo and Farmer 2004, and Lillo, Szabolcs, and Farmer 2005.

52．Szabolcs and Farmer 2005.

09

1．Dennett 1995.

2．道金斯和丹尼特指出，设计必须拥有设计师的直觉是神创论信仰的强大动力，缺乏对进化如何创造智能设计的理解是许多人仍然相信神创论或相信进化只是“一种理论”的重要原因。参见Dembski and Ruse, eds. 2004。

3．Dawkins 1986.

4．实验描述见Sims 1994a和1994b。

5．Dennett 1995.

6．从某种意义上来说，西姆斯的程序中有一个“创造者”，因为西姆斯设定了进化的条件。在自然界中，我们离能够讲述从大爆炸到生命起源再到今天的故事已经很近了。参见Fry 1999，Sole and Goodwin 2000，Morowitz 2002，Schneider and Sagan 2005，and Lane 2005。

7．Dennett 1995, pp. 48-60.

8．乐高是乐高集团的注册商标。

9．Kauffman 2000, pp. 223-225.

10．这个例子的灵感来自丹尼特的巴别塔图书馆。道金斯使用了一个类似的例子，他称之为贝壳博物馆，参见Dawkins 1996, pp. 180-203。

11．Holland 1975 and 1995, Mitchell 1996.

12．Hull 1988.

13．赫伯特·西蒙在1962年的一篇论文中讨论了等级制度和模块化在复

杂系统中的作用，参见Simon 1996, pp. 183-216。

14． Dawkins 1976 and 1982.

15．原则上，这种思想实验可以在现实生活中实现，人们会期望借此开发出有趣的乐高玩具。唯一的障碍是进化确实需要很多的迭代和很长的时间，这可能超出了两个7岁孩子的忍耐极限。这就是为什么这类实验通常要在计算机上模拟，因为计算机不容易疲劳。

16．From Dawkins 1976, pp. 15-20, Wilson 1992 and Fortey 1997.

17．相关讨论和进一步的参考资料参见Fry 2000，Morowitz 2002，Schneider and Sagan 2005。

18．Dawkins 1976.

19．Dawkins1976.

20．Kauffman 1995, p. 163.

21．Dennett 1995, pp. 111-113.

22．设计空间大小由模式的长度定义。尽管模式的长度没有概念上的限制，但是在任何运行的进化系统中，模式的长度将受到模式阅读器工作方式的限制。无限长的模式是不可能的，因为任何模式阅读器都无法在有限的时间内读取它们。

23．Dennett 1995, pp. 77-80.

24．在现实中，这样的适合度景观将有数十亿个维度。然而，为了便于讨论，我们将假设它只有三个维度。关于适合度景观维度的讨论参见Kauffman 1993、1995。我们还将假设小于最大100亿个碱基对的编码用“0”填充在空格中，因此一次移动一个符号也意味着向更长的或更短的基因组方向移动。

25．1931年，休厄尔·赖特率先讨论了适合度景观，考夫曼对其数学性质进行了广泛的探索。针对普通读者的精彩讨论参见Kauffman 1995 and Dennett 1995。

26． Kauffman 1995, pp. 161-189.

27．Crutchfield, pp. 101-133 in Crutchfield and Schuster, eds. 2003。

28．实际上有很多种爬山的算法，其中一种叫做随机突变。其他参见Mitchell 1996, p. 129 for others。

29．考夫曼在他的适合度景观模型中讨论了远距离跳跃和近距离跳跃混合的效果。

30．Kauffman 1995, p. 180.

31．Crutchfield, pp. 101-133 in Crutchfield and Schuster, eds. 2003.

32．关于双臂赌博机问题的讨论请参见Holland 1975，1995。关于组织范围内的探索与开发之间的紧张关系参见March 1991。关于在进化的背景下进行进一步的讨论参见Mitchell 1996, pp. 117-118，以及Axelrod and Cohen 1999, pp. 43-50。

33．关于双臂赌博机的证据和讨论参见Mitchell 1996, pp. 119-125。

34．Dennett 1995, pp. 77-78.

35．Dennett 1995, pp.128-135.

36．关于不同搜索算法适用于不同的景观结构有着丰富的文献：Whidey 1993、Kauffman 1993和Mitchell 1996。我并不是说进化是所有问题的最佳搜索算法，只是说进化在搜索动态的、粗相关的景观时具有许多有用的特征。大自然提供了一种存在证明，即进化是解决这类搜索问题的好方法。尽管目前还不存在，但我们还是有可能找到一种更优越的算法。

37．这些结果都不是预先确定的。进化并不一定带来秩序或增长，除非存在正确的参数。考夫曼认为生物进化会自我调整适应这些参数：Kaufman 1993。

38．Koza 1992.

39．Hodgson 1993.

40．Crutchfield and Schuster, eds. 2003, Koza 1992, and Mitchell 1996.

41．Henry IV, Part I, II, iv, line 460.

10

1．Axelrod 1997a, p. xi.

2．Fudenberg and Tirole 1991 and Gintis 2000. Brandenburger et. al. 1997.

3．Bowles 2004, p. 24.

4．合作在单轮博弈的案例中占有强大的主导地位。反向归纳法表明，在重复有限次数的博弈中，两名玩家都背叛是一种完美的子博弈平衡。在无限次数的博弈中，总是背叛的策略仍是一种完美的子博弈平衡。如果未来的贴现足够高， “在第一轮中合作，然后只要对手没有背叛就继续合作，一旦对手背叛，就在接下来的博弈中也选择背叛”也是一种完美的子博弈。在未知次数的有限博弈中，结果是不确定的，参见Fudenberg and Tirole 1991, pp. 110-111。

5．Axelrod 1984.

6．Axelrod 1997a, pp. 14-29.

7．Lindgren and Nordahl 1994, Lindgren in Arthur et. al. eds. 1997, pp. 337-367 and Lindgren and Johansson in Crutchfield and Schuster, eds. 2003, pp. 341-360.

8．生命游戏是细胞自动机最著名的例子之一，参见Wuensche and Lesser 1992 and Wolfram 1994 and 2002。

9．关系被一个随机数打破。格子边缘是连接在一起的，因此格子实际上是一个环面，而不是平面。

10．在这些阶段，行为者采取的是稳定的进化策略，相当于纳什均衡在生物学中的对应物。参见Maynard Smith 1982。

11．Dennett 1995, pp. 107-111.

12．为了了解丹尼特的把戏，可以把他的套装书策略发挥到极致。人们可以说有一个字母表图书馆，里面只有100本书，每本书里都只有一个字符。即使这个图书馆非常小，它仍包含着每一本可能的书。例如《白鲸》就是一部特别的1 250 000册的套书。同样，我们可以拥有一个比巴别图书馆更大的图书馆，里面有各种可能的5 000页的书籍。

13．如前所述，这是因为概要阅读器要先读取代码然后呈现设计，阅读器必须在有限的时间内读取代码。

14．Hannan and Freeman 1977.

15．Penrose 1959.

16．Nelson 2003.

17．Gould 2002.

11

1．Dawkins 1982, p. 109.

2．Carruthers & Chamberlain 2000. Cavalli-Sforza 2001.

3．当然，像大多数文化创新一样，贸易的“发现”可能不是一组人的单一行为。相反，这样一个有用的生存技巧可能是由多组原始人多次分别发现的，其中一组是智人的祖先。

4．Horan, Bulte, and Shogren 2005.

5．Ziman 2000. p. 3. Petroski 1992. Baldwin & Clark 2000.

6．Mokyr 1990, pp. 273-299, Mokyr in Chapter 5 of Ziman, ed. 2000, pp. 52-65, and Mokyr 2000.

7．莫基尔将他的设计空间定义为一组可行的技术和一组有用的知识。我对PT空间的定义与λ相对应。莫基尔的Ω就是PT进化所寻找的，并由模式阅读器呈现为物理存在。

8．Ridley 2001.

9．Cavalli-Sforza 2001, pp. 92-96, and Jones et al. 1992, pp. 107-143.

10．Anderson, Arrow, and Pines, eds. 1988, pp. 125-146.

11．Kauffman 1993, pp. 395-402, and 2000, pp.222-232.

12．Braha and Maimon 1998, pp. 6-7. Henderson and Clark 1990, and Baldwin and Clark 2000.

13．Baldwin and Clark 200. Braha and Maimon 1998, pp. 109-142.

14．Jones, et. al. 1992, pp. 350-360.

15．Kauffman 1995, pp. 191-206 and Kauffman 1995b.

16．Kauffman 1995, pp. 191-206 and Kauffman 1995b.

17．Ziman in Ziman, ed. 2000, pp. 3-12, and pp. 41-51.

18．Campbell 1960. Simon 1996, pp. 51-83. Aldrich 1999, pp. 22-26. Braha and Maimon 1998,pp. 19-84. Stuart Kauffman 1995, p. 202.

19 . Braha and Maimon 1998, p. 25.

20 . 有关便利贴发展的故事可以在3M公司的网站上找到。

21 . Petroski 1992, p. 22.

22 . 许多人认为这个故事是一个都市传奇或网络神话。1982年7月3日，飞机起飞后《纽约时报》和美国有线电视新闻网报道了这个故事。1992年12月19日，沃尔特斯因意外事故被联邦航空管理局处以罚款处罚。沃尔特斯在1993年自杀，随后，一部以他的生活和飞行梦想为原型的戏剧被创作出来。

23 . This is again building on the ideas of Simon 1996.

24 . Mokyr 1990; Mokyr, in Ziman 2000, pp. 52-65; and Mokyr 2000.

25 . Campbell and Pryce 2003.

26 . Foster 1986.

27 . Kauffman 1995a, pp. 202-206.

28 . Henderson and Clark 1990.

29 . Henderson and Clark 1990.

30 . Jardine 1999, Dear 2001 and Gribbin 2002.

12

1 . Easterly and Levine 2002.

2 . Lewis et al. 2002 and Johnson 2002. “US Productivity Growth 1995-2000”. Dorgan and Dowdy 2002.

3 . Nelson 2003.

4 . North 1990, p. 3. Scott 2001.

5 . Simon, 1996, pp. 139-167. Nelson and Winter, 1982. Loasby 1999, Hodgson, ed. 2002, and Bowles 2004.

6 . Krames 2002.

7 . Wright 2000.

8 . Diamond 1997, pp. 83-193.

9 . Freeman and Soete 1997，pp. 141-148.

10. Ibid., pp. 137-138.

11. Chandler 1962.

12. Axelrod 1984, Skyrms 1996, Axelrod 1997a, Young 1998, Gintis 2000, and Bowles 2004.

13. Wright 2000.

14. Wright 1994, p. 194.

15. Sylvia Nasar 1998.

16. Sober and Wilson 1998. Seabright 2004. Dawkins 1976 and Maynard Smith 1982.

17. Lindgren and Nordahl 1994 and Arthur et al. 1997, pp. 337-367.

18. Kagel and Roth 1995, pp. 253-348. Henrich, et. al. 2004.

19. A Wright 1994, Pinker 1997 and Pinker 2002.

20. Fukuyama 1995 and Harrison and Huntington 2000. Maynard Smith 1982 and Axelrod 1997b.

21. Sober and Wilson 1998.

22. Wright 1994, pp. 155-180.

23. Wright 2000, pp. 30-31.

24. Diamond 1997.

25. Seabright 2004.

26. Wright 2000, pp. 33-35.

27. Wright 1994.

28. Wright 2000, p. 33.

29. Wright 2000, pp. 39-41.

30. Easterly and Levine 2002.

31. 有关语言和参考文献演变的讨论，见第11章注释2。

32. Simon, 1992 and 1996. Cyert and March 1992. Prietula et al. 1998, Kennedy and Eberhart 2001, and Monge and Contractor 2003.

33 . Hutchins 1996.

34 . Jay 2000, p. 38.

35 . Ibid., pp. 121-122.

36 . Micklethwait and Wooldridge 2003.

13

1 . 我只是说基因是生物学中选择的单位，这仍然是一个活跃的研究领域，参见Dawkins 1976 and 1982。

2 . Nelson 1995, Hodgson 2002, and Knudsen 2002.

3 . Dawkins 1976 and 1982.

4 . 最佳实践管理研究和商学院案例证明，人类有回顾性地识别模块的能力。有关最佳实践研究的示例，请参阅公司执行委员会的网站和哈佛大学商学院的案例库。

5 . Nelson and Winter 1982, p. 14. Cohen and Sproull 1996. Dawkins 1976, p. 192. Dawkins 1982, Dennett 1995, Lynch 1996, Brodie 1996, Blackmore 1999, Aunger 2000, and Sherman 2002. Boyd and Richerson 1985, Dunbar, Knight, and Power 1999, Balkin 1998, Klein and Edgar 2002, Plotkin 2002, Boyd and Richerson 2005.

6 . 亚马逊网站上关于如何写商业计划的书竟然有97　041本。有研究表明，这些作者都遵循着相同的基本公式。我也可以证明这个公式的适用性，因为我曾经在一家风险投资公司工作过，在几年的时间里阅读过几百份商业计划书。

7 . Collins and Porras 1994, pp. 140-168.

8 . Niehans 1990, p. 447.

9 . Hayek 1948, 1988.

10 . Wright 2000.

11 . Cameron and Neal 2003. McMillan 2002. Bevir and Trentmann 2004.

12 . Baumol 2002.

13 . Chandler 1977.

14 . My use of the term resources in this context is consistent with its meaning in the management literature. Wernerfelt 1984 and 1995 and Robert Grant, pp. 179-199, in Segal-Horn 1998 for a review and references.

15 . Nelson & Winter 1982, pp. 119-121, and Nelson 1995, p. 69.

16 . Nelson 1995.

17 . Aghion and Howitt1998. Baumol 2002.

18 . Gintis 2003.

19 . Axtell 2002 and 2003.

20 . Baumol 2002.

21 . Kahneman, Diener, and Schwarz 1999 and Schwartz 2004.

22 . V Havel, *Summer Meditations* New York: Alfred A. Knopf, 1992, p. 62. McMillan p. 7.

23 . Campbell 1960. Plotkin 1993, 2000. Hull 1988.

24 . Cameron and Neal 2003, pp. 154-155.

25 . Johnson 1997.

26 . Ibid., pp. 94-95.

27 . Ibid., p. 94.

14

1 . 传记细节来自纽约大学创办的“经济思想史”网站。

2 . Cambridge, MA: Harvard University Press, 1966.

3 . Georgescu-Roegen 1971.

4 . Teilhard de Chardin 1969. Wright 2000.

5 . Hayek 1960 and 1988, and Boulding 1978.

6 . Schneider and Sagan 2005.

7 . Kauffman 1995a, pp. 44-45.

8．Georgescu-Roegen 1971, p. 18.

9．Roegen 1971, Daly 1999, and Mirowski 1989.

10．Daly 1999, pp. 78-79.

11．Ecological Economics 22, no. 3 September 1997: 271-273.

12．Georgescu-Roegen 1971, p. 21.

13．Georgescu-Roegen,1971, p.17. Mirowski, 1989, p. 367.

14．Mirowski 1989, p. 382.

15．Mirowski 1989, p. 383.

16．对于现代人来说，乔治斯库-罗根的一些论点看起来很古怪，甚至是错误的。他对熵的看法并没有吸收20世纪80年代和90年代信息理论研究的成果，他的进化观是在计算机能够对进化理论进行数学探索等的基础上发展起来的。尽管如此，他的总体思路仍然很好。

17．A. S. Eddington, *The Nature of the Physical World New York*: MacMillan, 1930, p. 74.Haynie 2001, p. xii.

18．Georgescu-Roegen 1971, pp. 6, 11, 196-198, 278.

19．Georgescu-Roegen 1971, p. 277.

20．Georgescu-Roegen 1971, p. 18.

21．Keizer 1987.

22．唯一能告诉我们时间的方向的方法是有一个参考点指向一组固定的恒星。但只有当我们事先知道“正确”的轨道是什么时，这些信息才会有用。我们可以说这个过程是可逆的，因为一个对太阳系一无所知的太空物理学家无法区分前向轨道和后向轨道，而任何熟悉热力学第二定律的外星人都能从一段打碎牛奶瓶的影片中分辨出时间的方向。

23．Prigogine 1996 and Deutsch 1997.

24．Christian von Baeyer 1998, p. 133.

25．Christian von Baeyer 1998.

26．在传统经济学中，可逆性问题没有得到明确的论述。然而，从可逆性的角度看传统理论，我们会发现一些有趣的矛盾。例如，交易通

常是不可逆的，例如瓦尔拉斯拍卖。然而，没有什么固有的东西可以阻止一个人向后运行新古典主义的生产函数。因此，新古典主义理论之所以有“时间之箭”（就其本身而言），是因为公共事业的安排，或者是最初的捐赠，或者是外生的冲击，使系统失去了平衡，使得公共事业通过不可逆的交易及时拉动经济向前发展。有趣的是，这种微观经济不可逆性并不会转化为宏观经济不可逆性。大多数宏观模型要么是时间可逆的，例如索洛的增长模型；要么是时间之箭外生的，例如钻石叠代模型。这进一步证明了传统微观和宏观理论世界之间缺乏一致性和整合性。

27．Christian von Baeyer 1998, pp. 152-155.

28．Smith and Foley 2002.

29．Christian von Baeyer 1998, pp. 146-152.

30．有人可能会提出这样的问题：破坏公物的人是为了享乐才破坏东西的。我认为，尽管这些人可能满足了一些心理需求，但他们并没有创造经济价值。而纳税人需要付钱给警察来阻止这些人，从社会的角度来看，破坏公物是一种破坏价值的行为。

31．即使我们将拆除这一中间步骤隔离开来，并将其本身视为一个转变，我们也会看到，虽然实体建筑的熵在增加，但随着拆除公司不断扩大，雇用人员、建立组织、购买设备、创造关于拆除的知识等，它的熵是在减少的。整个经济系统中局部的熵减小，随着建筑物的瓦砾回到环境中，宇宙的熵增大了。

32．感谢2002年密歇根大学-圣塔菲研究所复杂性研讨会的参与者提出了这个例子。

33．Christian and Baeyer 1998, pp. 157-158.

34．Bowles 2004, pp. 93-126. Buss, 1999.

35．Wright 1994, Pinker 1997 and 2002, Ridley 1997 and Buss 1999.

36．Wright 1994.

37．Consumer Expenditures in 2000, *U.S. Bureau of Labor Statistics*, April 2002.

38．Abul-Magd 2002.

39．Pinker 1997, pp. 374-378.

40 . Pinker 1997, pp. 378-385.

41 . Wright 1994.

42 . Fox 2002.

43 . www.telenor.no/fou/program/nomadiske/artikler.shtml. Accessed on August 15, 2004.

44 . Pinker 1997, p. 522.

45 . Pinker 1997, pp. 521-538.

46 . Laland, Odling-Smee, and Feldman 2000 and Bowles and Gintis 2002.

47 . Klein and Edgar 2002, pp. 194-195. Pinker 1997, pp. 528-538.

48 . Georgescu-Roegen 1971, pp. 283-291.

49 . Dawkins 1976, 1998, Dennett 1995, and Wright 2000.

50 . Kahneman et al. 1999 and Layard 2005.

51 . Wright 2000. Horan, Bulte, and Shogren 2005.

52 . Cohen and Sproull, eds. 1996, Mintzberg, Ahlstrand, and Lampel 1998, pp. 175-231, and Eisenhardt and Santos in Pettigrew, Thomas, and Whittington, eds. 2002, pp. 139-164.

53 . Plotkin 1993.

54 . Gould 2002.

55 . Hausman, ed. 1994, pp. 10-15 and Cartwright 1999, pp. 49-74.

15

1 . Mas-Colell, Whinston, and Green 1995, pp. 167-215.

2 . “And Now, the War Forecast,” *The Economist Technology Quarterly*, September 17, 2005,pp. 21-22.

3 . Flake 1998, Weinberger 1996.

4 . Mintzberg, Ahlstrand, and Lampel 1998 and Pettigrew, Thomas, and Whittington eds. 2002.

5 . Chandler 1962, p. 13.

6. Mintzberg, Ahlstrand and Lampel 1998. Rumelt, Schendel and Teece, eds. 1994 and Pettigrew, Thomas, and Whittington eds. 2002.

7. Pettigrew, Thomas, and Whittington eds. 2002, pp. 55-71.

8. Pettigrew, Thomas, and Whittington, eds. 2002, pp. 55-71.

9. Ghemawat 1991.

10. www.ormiston.com/annieoakley accessed August 15, 2004.

11. Gell-Mann 1994, pp. 133-134.

12. www.yogi-berra.com.

13. Van der Hijden 1996 and Schwartz 1991. Courtney 2001.

14. Mas-Colell et al. 1995.

15. Holland et al., 1989.

16. Lee et al. 1998, Axtell 2001, and Powell 2003. Camerer et al. 2004, pp. 9-12.

17. Foster and Kaplan 2001, ch. 3.

18. Micklethwait and Wooldridge 2003.

19. Wiggins and Ruefli 2005. Wiggins and Ruefli 2002.

20. Foster and Kaplan 2001.

21. 在合并或兼并中，即便被并购公司的所有权改变了，其资产也是存在的。但我认为这属于“放弃清单”，因为在第三部分的概述框架中，资产正处于另一个业务计划的控制之下。并购是成功的商业计划加强对资源控制的一种机制。

22. Wiggins and Ruefli 2002.

23. Wiggins and Ruefli 2002 and 2005.

24. 绩效是通过资产回报率和托宾的q来衡量的，所述结果是资产回报率。

25. Powell 2002.

26. 参见鲁弗利2005年9月25—26日与作者的往来邮件。

27 . “A new evolutionary law,” *Evolutionary Theory*, vol. 1, pp. 1-30. Kauffman 1995b.

28 . Henderson and Mitchell, 1997, Pettigrew, Thomas, and Whittington, eds. 2002, and Hoopes, Madsen, and Walker, 2003. Mintzberg, Ahlstrand, and Lampel 1998.

29 . 参见鲁弗利2005年9月25—26日与作者的往来邮件。

30 . Hannan and Freeman 1977, 1984, 1989, and Hannan and Carroll 1992.

31 . Foster 1986, p. 132-135 and Tushman and O’Reilly 1997, pp. 17-25.

32 . Hannan and Freeman 1989 and Baum and Singh 1994.

33 . Foster and Kaplan 2001, p. 20.

34 . Robert Boyd and Peter Richerson 1985, Nelson and Winter 1982, Donald Campbell 1960,Burgelman 1983, Barnett and Burgelman 1996, and Burgelman’s 2002, Peters and Waterman 1982, p. 114 and Hamel 2000, pp. 264-269, and pp. 297-306.

35 . Avinash Dixit and Robert Pindyck 1994, Beinhocker 1999, Bryan 2002.

36 . Beinhocker 1999.

37 . Beinhocker and Kaplan 2002.

38 . Foster and Kaplan 2001, Chapter 11.

39 . Mintzberg, Ahlstrand, and Lampel 1998, pp. 176-231.

40 . Dixit and Pindyck 1994, Copeland and Keenan 1998, and Leslie and Michaels 2000.

41 . Bryan 2002.

42 . Peters and Waterman 1982, March 1991, Collins and Porras 1994, Tushman and O’Reilly 1997, and Foster and Kaplan 2001.

43 . Bryan 2002.

44 . Wiggins and Ruefli 2002, p. 95.

45 . Collins and Porras 1994, Tushman and O'Reilly 1997, and Lencioni 2002.

46 . Collins and Porras 1994, p. 94.

47 . Collins and Porras 1994, p. 95.

48 . Welch 2001.

49 . Camazine et al. 2001, pp. 188-215.

50 . J. Doneheyand G. Overholser of CapitalOne at the Ernst & Young Embracing Complexity Conference, Boston, M.A., August 2-4, 1998 Bonabeau and Meyer 2001.

51 . Kaplan et al. 1996.

52 . Campbell et al. 2003.

53 . Hamel 2000.

54 . Berzins et al. 1998, p. 7.

16

1 . www.westinghousenuclear.com and www.ge.com, and Collins and Porras 1994, Slater 1999,and Welch 2001.

2 . Welch 2001 and Collins and Porras 1994.

3 . Collins and Porras 1994.

4 . Slater 1999 and Welch 2001.

5 . Krames 2002 and Slater 1999.

6 . Arrow 1974, Simon 1997, Williamson and Winter 1991, Jensen 1998, Hatch 1997, Penrose 1959, Williamson 1990, Cyert and March 1963, Astely et al. 1983, Koza and Thoenig 1995,and Giddens 2001. Hatch 1997.

7 . Nelson and Winter 1982, Aldrich 1999, Dosi, Nelson, and Winter 2000, Haken 2000, Lee et al. 1998, Lesourne and Orléan 1998, Morecroft and Sterman 1994, Prietula et al. 1998,Axelrod 1984 and 1997a, Axelrod and Cohen 1999, Durlauf and Young 2001, Young 1998,and Monge and Contractor 2003.

8 . Aldrich 1999.

9 . Cyert and March 1992, Williamson, ed. 1995.

10 . Aldrich 1999, p. 4.

11 . Roberts 2004, pp. 74-117.

12 . Williamson and Winter 1993.

13 . Ibid, and Williamson 1995.

14 . Holmstrôm and Roberts 1998.

15 . Camazine et al. 2001, pp. 23-26.

16 . Cohen and Sproull, 1996, pp. 516-540, Hutchins 1996.

17 . Wright 2000.

18 . Axelrod and Cohen 1999, pp. 43-50.

19 . Peters and Waterman 1982.

20 . Collins and Porras 1994, Foster and Kaplan 2001, and Tushman and O'Reilly 1997.

21 . March 1991.

22 . "Help At Last," *The Economist*, November 29, 2003, p. 11.

23 . Cohen and Sproull 1996 and Camerer, Loewenstein, and Rabin 2004.

24 . Bob Woodward's Bush at War New York: Simon & Schuster, 2002 and Richard Clarke's Against All Enemies New York: Free Press, 2004.

25 . Lovallo and Kahneman 2003.

26 . Ibid., p. 5.

27 . Welch 2001, p. 106.

28 . Holland et al. 1986 and Holland 1995.

29 . Sternberg 1999, pp. 122-123, Holland et al. 1986, pp. 250-254.

30 . Margaret Thatcher, "Special Report: Politics Past," politics.guardian.co.uk.

31 . Harrington 1998.

32 . Krames 2002, p. 105.

33 . Page 1996.

34 . Penrose 1959.

35 . Wernerfelt 1984 and 1995, Barney 1991, Henderson and Mitchell 1997, Segal-Horn 1998,Mintzberg, Ahlstrand, and Lampel 1998, and Dosi, Nelson, and Winter 2000.

36 . Penrose 1959.

37 . Prahalad and Hamel 1990, Baghai, Coley and White 1999.

38 . Prietula, Carley and Gasser 1998, pp. 3-22.

39 . Boyd and Richerson 1985, p. 33.

40 . Balkin 1998, Aunger 2000, and Sherman 2002. Dunbar et al. 1999 and Klein and Edgar 2002.

41 . Nelson and Winter, 1982.

42 . Katz 2000, Wright 1994, Ridley 1996, and Pinker 2002.

43 . Henrich et al. 2004.

44 . Anna Karenina, part I, ch. 1. Deal and Kennedy 1982, Peters and Waterman 1982, Kotter and Heskett 1992, Schein 1992, Howard and Haas 1993, Collins and Porras 1994, Cohen and Sproull 1996, de Gues 1997, Tushman and O'Reilly 1997, Aldrich 1999, Cohen and Prusak 2001, Collins 2001, Welch 2001, Bower 2003, and Roberts 2004.

45 . Collins and Porras 1994, p. 147.

46 . Edersheim 2004. www.mckinsey.com.

47 . Krames 2002.

48 . Gardner 2004.

49 . Krames 2002.

50 . Collins and Porras 1994, Collins 2001, Slater 1999, Foster and Kaplan 2001, Tushman and O'Reilly 1997, and Roberts 2004.

51 . Penrose 1959.

52 . Ibid., p. 18.

53 . Axtell 2001.

54 . Penrose 1959, pp. xvi-xviii, 12-13,18-19, 213-214.

55 . Axtell 2001, Stanley et al. 1996, and Amaral et al. 1997 and 1998.

56 . Minsky 1985.

17

1 . Lewis 1999 and Lowenstein 2000.

2 . Lewis 1999.

3 . Peter Bernstein 1992.

4 . Mandelbrot, 1997, p. 24.

5 . Malkiel 1973.

6 . Brealey and Myers 1988, Cochrane 2005, and Ross 2005, Lengwiler 2004

7 . Copeland et al. 1994.

8 . Campbell et al. 1997 p. 20.

9 . Bernstein 1992, p. 17.

10 . Fischer Black, “Noise,” in Thaler 1993, pp. 3-22.

11 . Malkiel 1973, pp. 161-172.

12 . Thaler 1993, Shleifer 2000, Montier 2002, and Camerer, Loewenstein, and Rabin 2004, Ross 2005.

13 . Gleick 1987, pp. 83-90.

14 . Mandelbrot 1997, pp. 371-411.

15 . 更具体地说，曼德尔布罗特提出数据遵循列维稳定分布（Levy Stable Distribution），尾部具有无标度的特点。

16 . Mandelbrot and Hudson 2004.

17 . Mandelbrot and Hudson (2004), p. 166.

18 . Mandelbrot 1997, pp. 444-457.

19 . Mandelbrot 1997, pp. 419-443.

20 . Lo and MacKinlay 1999, p. 4.

21 . Fama 1970, Lo and MacKinlay 1999, p. 13.

22 . Lo and MacKinlay 1999, p. 14

23 . Lo and MacKinlay 1999, pp. 17-45, Mandelbrot, 1997, and Mandelbrot and Hudson, 2004 Engle 1982.

24 . LeBaron 1989 and 1994, Campbell et al. 1997, and Lo and MacKinlay 1999.

25 . Farmer 1999, Farmer and Lo 1999, Mantegna and Stanley 2000, Sornette 2003, and Johnson, Jeffries, and Hui 2003.

26 . Mantegna and Stanley 1996, Plerou et al. 1999, and Cont and Bouchaud 2000, and Sornette 2003.

27 . Gleick 1987 and Bass 1999.

28 . Farmer 2001.

29 . Arthur 1995, Arthur et al. 1997, p. 38.

30 . Arthur et al., "Asset Pricing Under Endogenous Expectations in an Artificial Stock Market,"in Arthur, Durlauf, and Lane 1997, pp. 15-43.

31 . Kahneman et al. 1982, Thaler 1993, Kagel et al. 1995, and Shleifer 2000.

32 . LeBaron et al. 1999.

33 . Farmer 1998, p. 5.

34 . Farmer 1998 and 2002.

35 . Farmer 2001, p. 64.

36 . O'Hara 1995.

37 . Farmer 1998, 2001, and 2002.

38 . Neftci 1991, Brock et al. 1992, and LeBaron 1998.

39 . Campbell et al. 1997, pp. 253-287.

40 . Bernard, "Stock Price Reactions to Earnings Announcements," in Thaler 1993, pp. 303-340.

41 . "A Survey of Risk" , *Economist*, January 24, 2004, pp.10-11.

42 . Farmer 1998, p. 4.

43 . Surowiecki 2004.

44 . Surowiecki 2003.

45 . *General Theory of Employment, Interest and Money*.

46 . Farmer 1998, pp. 44-45, and Farmer 2002, p. 935.

47 . Lo 2004.

48 . Copeland, Koller, and Murrin 2000.

49 . Brealey and Myers 1988. Cochrane 2005.

50 . Campbell, Lo, and MacKinlay 1997, pp. 188-217, and Montier 2002, pp. 81-86.

51 . 1977年，理查德·罗尔发现了这个问题。市场投资组合通常代表的是价值加权的CRSP指数。

52 . Karceski 2002.

53 . Campbell, Lo, and MacKinlay 1997, pp. 211-212, and Monder 2002, pp. 83-84. Copeland, Roller, and Murrin 2000, pp. 224-226.

54 . 该理论认为比较仪应该是零beta组合，但必须再次使用代理，通常是政府债务作为代理，并且必须在选择代理时选择，如时间框架、特定工具和发行政府。

55 . Campbell, Lo, and MacKinlay 1997, pp. 219-251, Lo and MacKinlay 1999, pp. 189-212,and Montier 2002, pp. 87-92.

56 . Andrew Hill, reported by Financial Times, May 14, 2003.

57 . McNulty et al., 2002, and Shefrin, 2005.

58 . Copeland, Koller, and Murrin 2000, p. 7.

59 . De Swaan and Harper 2003.

60 . Kaplan and Norton 1996 and 2000.

61 . Dobbs and Koller 2005.

62 . Jensen 1998 and 2001, and Mas-Colell, Whinston, and Green 1995, pp. 152-154.

63 . Copeland, Koller, and Murrin 2000, pp. 3-15.

64 . Elkington 1999, Kelly 2003, Zadek 2004, and Bakan 2004.

65 . Handy 2002.

66 . Foster and Kaplan, 2001, p. 8.

67 . 公司也可以在没有外部投资的情况下从自身的现金流中获取增长，但此类收益不会返还给股东，而是作为资本重新投资于企业生产。

68 . Graham, Harvey, and Rajgopal 2005.

69 . Davis 2005a.

70 . Davis 2005b

71 . de Gues 1997, p. 11.

72 . Collins and Porras 1994, pp. 48-79.

18

1 . Johnson 1988, p. 4.

2 . Fukuyama 1992.

3 . Giddens 1998 and 2000.

4 . Giddens 2000.

5 . Gintis et al. 2005, p. 4.

6 . Gintis et al. 2005, p. 3.

7 . Ibid., p. 8.

8 . Gintis et al. 2005 and Henrich et al. 2004.

9 . Silk in Gintis et al. 2005, pp. 43-75.

10 . Damasio 2005.

11 . Gintis et al. 2005, p. 29, and Heinrich et al. 2004, pp. 125-167.

12 . Fong, Bowles, and Gintis, in Gintis et al. 2005, pp. 278-302.

13 . Arrow 1974.

14 . Wright 2000.

15 . 英国撤销电话目录的故事可以在英国广播公司2004年5月20日播出的《错误的号码》（Wrong Numbers）这一纪录片中找到详细介绍。

16 . Fligstein 2001.

17 . Morris 2001.

18 . www.worldbank.org.

19 . www.worldbank.org.

20 . Landes 1998, p. xix

21 . North and Thomas 1973, Rosenberg and Birdzell 1986, and Landes 1998.

22 . Diamond 1997.

23 . Harrison and Huntington 2000.

24 . Ibid., p. 2.

25 . Bednar and Page 2003 and Bowles 2001a.

26 . Grondona, in Harrison and Huntington 2000, pp. 44-55.

27 . Thomas and Mueller 2000.

28 . Harrison 2001.

29 . Hammond 2000.

30 . Zak and Knack 2001, Zak 2003, and Glimcher 2002 and 2003.

31 . Henrich et al. 2004.

32 . Zak 2003, Gintis et al. 2005, p. 381.

33 . Fukuyama 1995, pp. 62-145.

34 . Hoff and Sen 2001.

35 . Easterly 2002.

36 . Putnam 2000.

37 . Ibid., p. 19.

38 . Ibid., pp. 21-22

39 . Putnam 1993, p. 5.

40 . Ibid., pp. 5-6.

41 . Padgett and Ansell 1993.

42 . Putnam 2000, pp. 15-17.

43 . Purdy 2003, Fukuyama 1995.

44 . Fukuyama 1999, pp. 31-44.

45 . Putnam 2000, p. 195.

46 . Ibid., pp. 402-414.

47 . Fukuyama 1999, pp. 263-282.

48 . Putnam 1993 and 2000, Fukuyama 1995 and 1999, Bowles, Choi, and Hopfensitz 2003,Skyrms 1996, Sober and Wilson 1998, and Bowles 2004.

49 . Bowles 2004, p. 12.

50 . Stanley Fischer, "Global Inequality: More or Less Equal," The *Economist*, May 11, 2004.

51 . Abul-Magd 2002.

52 . Miller 2003, p. 50. Matt Miller.

53 . Arrow, Bowles, and Durlauf 2000, and Bowles, Gintis, and Osborne Groves 2005.

54 . Ibid., pp. 1-22.

55 . Duncan et al., in Bowles, Gintis, and Osborne Groves 2005, pp. 23-79. Bowles, Gintis, and Osborne 2001.

56 . Feldman, Otto, and Christiansen, in Arrow, Bowles, and Durlauf 2000, pp. 62-85.

57 . Bowles 2001b.

58 . Durlauf 2001.

59 . Miller 2003, pp. 69-91.

60 . Hertz, in Bowles, Gintis, and Osborne Groves, eds. 2005, p. 188.

61 . Harrison and Huntington 2000, p. xiv.

62 . Bewley, in Gintis et al. 2005, pp. 303-338.

63 . Wackernagel et al. 2002.

64 . Sterman and Sweeney 2002.

65 . Richards, ed. 2000, and Kollman, Miller, and Page, eds. 2003.

66 . *The Oxford Dictionary of Quotations Oxford*: Oxford University Press, 1999.

67 . E. M. Forster, 1938, *Two Cheers for Democracy*.

后记

1 . Fogel 2004 and Bernstein 2004.

2 . Diamond 2005.

3 . www.lib.umich.edu/govdocs/elec2004.

4 . Buder et al. 1997 and Johnson, Manyika, and Yee 2005.

5 . Huntington 1996.

6 . Kennedy 1987, Diamond 1997, and Wright 2000.

7 . Edmund Burke, 1790. Reflections on the Revolution in France.

译者后记

2018年6月，湛庐文化的简学老师和黄小骑老师向我推荐了他们花费数年时间、付出诸多努力而成功引进的《财富的起源》，并邀请我担任翻译。由于本人时间精力有限，于是我又邀请了刘玮琦、尤娜两位朋友与我共同翻译，我负责翻译的是1—9章的内容，刘玮琦负责的是10—14章的内容，尤娜负责的是15—18章以及后记的翻译。在此，我想感谢两位朋友的通力合作，感谢简院长与黄主编对我们工作的大力支持，使得本书得以顺利出版。

我在经济学方面的翻译经验可谓有限，平日里对于经济学读物也总怀有一种距离感。然而，在翻译本书的过程中，我深深地被书中的内容所吸引。本书为我们呈现了诸多事实，提供了一套创新性的、富有洞见的理论框架，为现代经济社会提供了更加贴切的解释和指导，让人耳目一新、深受启迪。相信读完此书的读者都会拥有这样的感受。由于我们的能力有限，译文中难免存在一些错误和不足之处，恳请广大读者多提宝贵建议和意见！我们也希望在未来能够翻译更多这类颇具价值的读物以飨读者。

俸绪娴

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读APP：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”APP。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读APP玩转指南

湛庐阅读APP结构图:



三步玩转湛庐阅读APP:



APP获取方式：
安卓用户前往各大应用市场、苹果用户前往APP Store
直接下载“湛庐阅读”APP，与最聪明的人共同进化！

使用APP扫一扫功能，
遇见书里书外更大的世界！



快速了解本书内容，
湛庐千册图书一键购买！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

延伸阅读

《复杂经济学》

◎布莱恩·阿瑟继《技术的本质》之后，又一部重磅作品！

◎圣塔菲研究所元老、斯坦福大学经济学教授、“复杂经济学”创始人、“拉格朗日奖”“熊彼特奖”得主布莱恩·阿瑟重磅新书。

◎经济学的新古典主义时代已经结束，取而代之的是复杂性时代！以独特的方式思考经济，把经济看作是不断自我计算、自我创建和自我更新的动态系统。





《多样性红利》

◎“得到”APP《精英日课》专栏作者万维钢，苇草智酷创始合伙人段永朝、北京师范大学系统科学学院教授狄增如、集智俱乐部创始人张江、阿里巴巴复杂科学研究中心副主任吕琳媛等联袂推荐！诺贝尔经济学奖得主肯尼斯·阿罗、著名心理学家菲利普·泰洛克等盛赞！

◎21世纪本就是各种各样知识大爆发的时代，必须建立跨学科思维范式，需要具备多样性认知工具箱：多样性视角、启发式、解释和预测模型。这些工具既是认知升级的阶梯，也是实现知识大融通的把手。





《为什么需要生物学思维》

◎应该怎么看待这个越来越复杂的世界？复杂系统研究专家塞缪尔·阿贝斯曼用《为什么需要生物学思维》这本书为我们提供了一个洞悉复杂世界的思考方式，告诉我们不必害怕。

◎北京大学国家发展研究院教授、财新传媒学术顾问汪丁丁，《连线》创始主编、畅销书《失控》作者凯文·凯利，美国经济学家、乔治梅森大学经济学教授泰勒·考恩，美国数学家、康奈尔大学数学教授、畅销书《同步》作者斯蒂芬·斯托加茨等联袂推荐！





《技术的本质》

◎复杂性科学奠基人、首屈一指的技术思想家、“熊彼特奖”得主布莱恩·阿瑟作品！

◎谷歌Java程序开发的灵感源泉，技术理论体系的先河之作，前所未有的关于技术产生和进化的系统性理论！

◎一次打开“技术黑箱”的尝试性创新探索，技术思想大师用平实的语言向大众讲述技术前沿思想！





(1)世界著名哲学家、认知科学家，美国艺术与科学院院士。全球50位最具影响力的健在哲学家之一，其哲学著作《直觉泵和其他思考工具》中文简体字版已由湛庐文化策划、浙江教育出版社出版。——编者注

(2)美国第十三任联邦储备委员会主席，美国国家经济政策的权威和决定性人物。讲述其传奇人生经历的《格林斯潘传》中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

(3)蒲式耳是谷物计量单位，一蒲式耳约为36升。——译者注

(4)这被视为美国跨学科领域的最高荣誉。——译者注

(5)美国行为科学家，因对决策过程的研究而著名。讲述其重量级研究成果的《特沃斯基精要》中文简体字版即将由湛庐文化策划、浙江人民出版社出版。——编者注

(6)这两个术语在经济学中有着另外的含义，指的是某个经济体是否参与国际贸易，但我们这里用的是它们在物理学中的定义。

(7)广派教是英国基督教圣公会的一个教派，支持对圣公会教条和仪式进行广泛、自由的解释，反对神学中的固有解释。——译者注

(8)整体具有而部分不具有的特性被称为涌现性。

(9)通用电器董事长兼首席执行官。——译者注

(10)“小世界网络之父”，网络科学奠基人之一，其颠覆式新作《反常识》中文简体字版已由湛庐文化策划，四川科学技术出版社出版。——编者注

(11)康奈尔大学应用数学教授，在混沌理论和复杂性理论方面的开创性研究工作获得了广泛认可。其最新力作《同步》中文简体字版已由湛庐文化策划、四川人民出版社出版。——编者注

(12)美国知名演员，代表作有《我爱迪克》《杀手信徒》等。——编者注

(13)计量单位，一蒲式耳约为36升。——编者注

(14)以第一个发现基因的19世纪僧侣的名字命名。

(15)哈佛大学心理学教授，世界著名思想家，其经典之作《思想本质》《心智探奇》《语言本能》《白板》《当下的启蒙》中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

CHEERS
湛庐



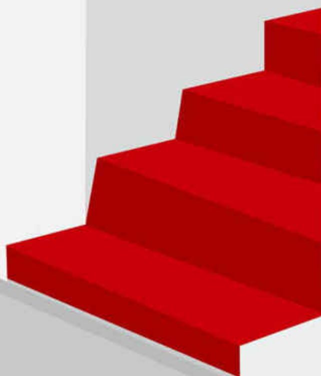
INTUITION PUMPS

and other tools for thinking

直觉泵

和其他思考工具

[美] 丹尼尔·丹尼特 ◎ 著
（Daniel C. Dennett）
冯文婧 傅金岳 徐韬 ◎ 译



浙江教育出版社
ZHEJIANG EDUCATION PUBLISHING HOUSE

版权信息

本书纸版由浙江教育出版社于2018年11月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：直觉泵和其他思考工具

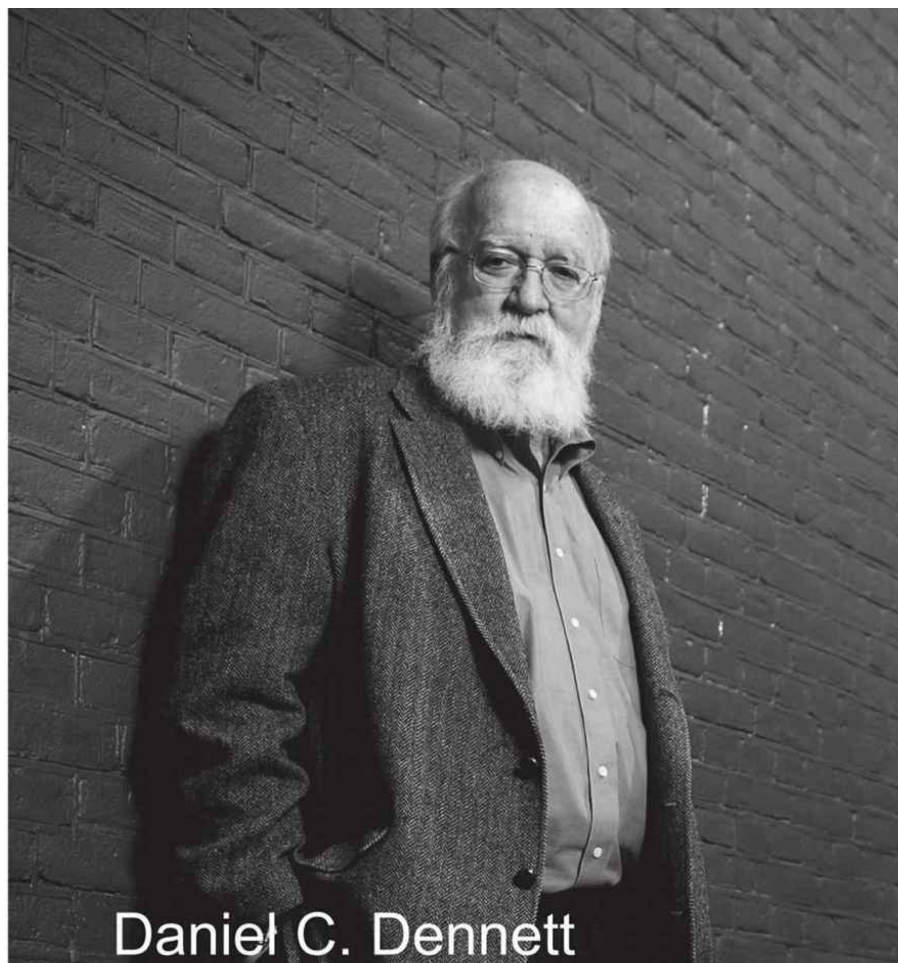
著者：丹尼尔·丹尼特

电子书定价：116.99

Intuition Pumps and Other Tools for Thinking by Daniel C. Dennett

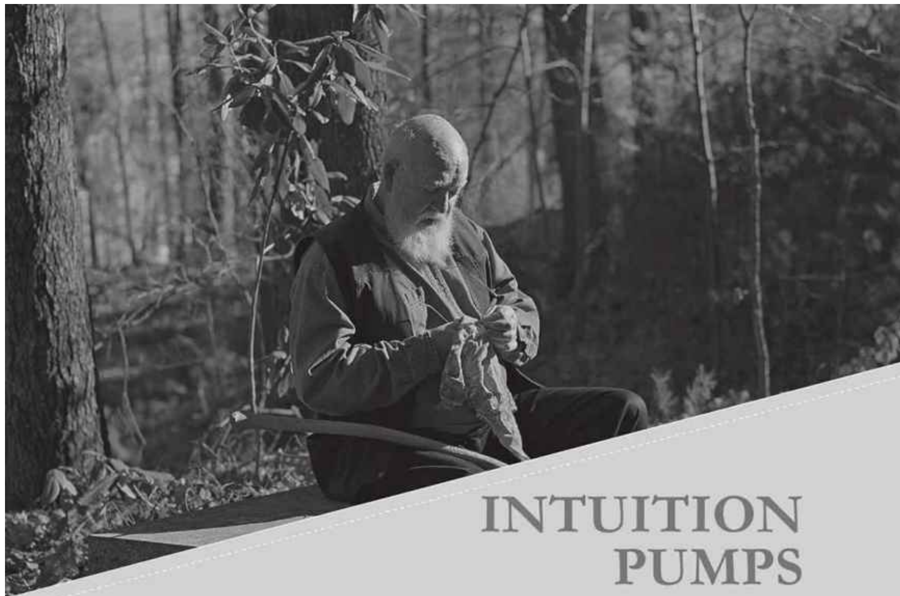
Copyright © 2013 by Daniel C. Dennett

丹尼尔·丹尼特



丹尼尔 · 丹尼特

著名美国哲学家、认知科学家



INTUITION PUMPS

AND OTHER TOOLS FOR THINKING

极具影响力的哲学家

1942 年，丹尼特出生于波士顿。他 11 岁时参加了一个夏令营，第一次接触到哲学。当时，一位辅导员告诉他：“丹尼尔，你知道你是谁吗？你是一名哲学家。”

17 岁时，丹尼特进入维思大学并对哲学有了兴趣。他发现了蒯因的著作《从逻辑的观点看》，想要与这位 20 世纪最重要的哲学家之一当面对质。于是，他大二就转学到了哈佛大学。3 年后，在蒯因的推荐下，他来到了牛津大学跟随吉尔伯特·赖尔攻读哲学博士学位。

目前，丹尼特是塔夫茨大学哲学教授，也是塔夫茨大学认知研究中心的掌门人。他曾获心智与脑奖、伊拉斯谟奖，两次获古根海姆奖。2001 年，丹尼特荣获被誉为“心灵哲学诺贝尔奖”的让·尼科奖。

丹尼特在心灵哲学方面有极其重要的贡献。他创造了“意向立场”“他者现象学”“直觉泵”等概念，论证意识的“困难问题”并不存在。美国教育网站 TheBestSchools 评选丹尼尔·丹尼特为全球 50 位最具影响力的健在哲学家之一。

INTUITION PUMPS

AND OTHER TOOLS FOR THINKING

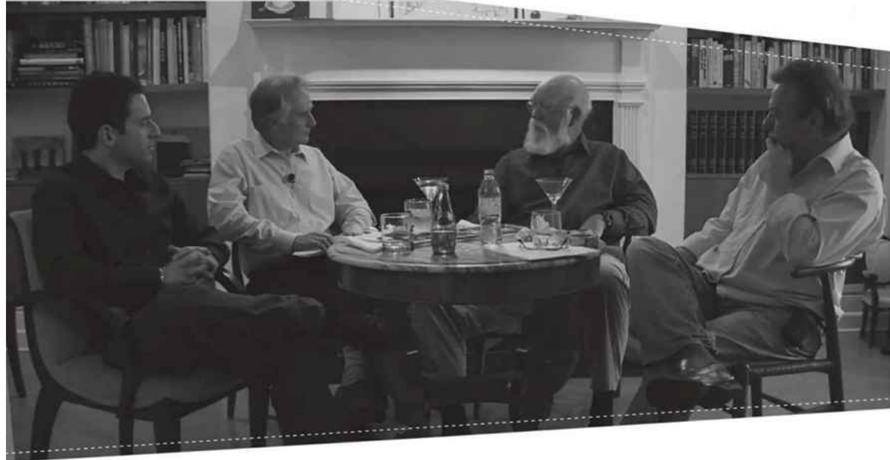
圣塔菲研究所米勒学者

圣塔菲研究所成立于 1984 年，它致力于从多学科角度研究复杂适应系统的基本原理，是世界闻名的复杂科学研究中心。2016 年，圣塔菲研究所名列“20 大世界顶级科技智库”。

这座“没有围墙”的学术圣地设置了大量外聘职位，其中最为尊崇的是米勒学者。该席位遴选出极少数杰出而具有创造力的思想家，为其提供独一无二的学术环境。

2010 年，丹尼尔 · 丹尼特当选圣塔菲研究所首届米勒学者，跨学科开展认知科学和进化生物学研究。随后的当选者还有与其丈夫史蒂芬 · 平克齐名的小说家、哲学家瑞贝卡 · 戈尔茨坦（Rebecca Goldstein），著名量子计算专家塞思 · 劳埃德（Seth Lloyd），科幻作家尼尔 · 斯蒂芬森（Neal Stephenson）等。





INTUITION PUMPS

AND OTHER
TOOLS FOR THINKING

新无神论四大骑士

丹尼特强烈支持进化论，将自然选择视作一种算法过程。他与宗教团体开展了大量的辩论，与理查德·道金斯（Richard Dawkins）、萨姆·哈里斯（Sam Harris）、克里斯托弗·希钦斯（Christopher Hitchens）并称新无神论四大骑士。

在《达尔文的危险观念》（*Darwin's Dangerous Idea*）和《打破魔咒》（*Breaking the Spell*）两本书中，他分别论证了道德和宗教可以起源于进化过程，在美国激起了强烈的反响。

丹尼特成长在一个公理会家庭，虽然他是一名无神论者，但他很喜欢去教堂。40 多年来，每年 12 月份，丹尼特都会在家里举办圣诞颂歌聚会，并充满深情地唱赞歌。

作者演讲洽谈，请联系
speech@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS 特别制作

赞誉

汪丁丁　北京大学国家发展研究院教授

2007年，朗润园十周年庆典，邀请了十位曾获得过诺贝尔奖的经济学家，其中一位是实验经济学家弗农·史密斯，需要我作陪。席间，我列举了一些跨界科学家的名字，希望听他评论。出乎意料的是，他只读过丹尼特的书，而且受益良多。我也喜欢丹尼特，并收藏了他的主要作品。《直觉泵和其他思考工具》以及2017年出版的*From Bacteria to Bach and Back: The Evolution of Minds*，尤其值得收藏。因为，这些是丹尼特50年悠闲思考的精华概述。前者试着推荐几十种不用数学但依赖隐喻的（批判性）思考工具给普通读者，后者则是对他毕生心灵哲学的一次总结，他很可能继续发表新的总结。

傅小兰　中国科学院心理研究所所长，中国科学院大学心理学系主任，中国心理学会理事长

本书旨在提供一本“人人都能读懂的书”，使人们掌握思考思想本身的有效工具。作者如数家珍地介绍了“归谬法”“拉波波特法则”和“史特金定律”等12种通用思考工具，又依次介绍了关于意义和内容、计算机、进化论、意识和自由意志等的思考工具。作者兼具哲学家和认知科学家的理论素养，也有很深的文学功底，尤其擅长使用比喻，使得本书读起来十分生动有趣。我愿将本书推荐给大家阅读参考。

周晓林　北京大学脑与认知科学中心主任，浙江师范大学心理与脑科学研究院院长，教育部长江学者特聘教授

丹尼尔·丹尼特将人类思想中最闪耀的星星摘下来，放进此书，送给读者。他通过使用思想实验和思维工具，将深奥的哲学问题阐述得浅显易懂，即便是只接受过基本文化教育的读者，也能跟随此书思考有关心智、生命和自由意志的问题。读者甚至可以把它作为一本工具书，通过学习书中提到的诸多思想、逻辑工具，训练在科学研究、辩论和谈判等思想过程中的“拳脚功夫”。

段永朝　苇草智酷创始合伙人，财讯传媒首席战略官

《直觉泵和其他思考工具》这本书读起来很费脑，要反复看很多遍，才能领会全书77个思想实验的精妙之处。丹尼尔·丹尼特是一位个性鲜明、富有科学家精神的著名哲学家，他的书，应当全力推荐！

王　烁　财新传媒总编辑

丹尼尔·丹尼特是现代进化论科普大家，也是最具科学家气质的哲学家之一。他首创的“直觉泵”这个概念告诉我们：在没有证据、数据的地方，我们可以用故事说话，从公认的直觉开始做思想实验，看它将我们带到何方。我强烈推荐《直觉泵和其他思考工具》这本书！

苏德超　武汉大学哲学学院教授，博士生导师

通过把当代科学的底层观念和传统哲学的深刻问题结合起来，丹尼特正在改变我们的直觉。他精通多个知识门类，致力于基础观念的澄清和革新，他比任何人都更接近传统意义上的伟大哲学家。

理查德·道金斯　著名生物学家，畅销书《自私的基因》作者

这是我读过的最好的一本书。

马文·明斯基　“人工智能之父”，图灵奖得主

丹尼尔·丹尼特是我们这个时代最杰出的哲学家之一，他就是下一个伯特兰·罗素。与传统哲学家不同，丹尼特还精通神经科学、语言学、人工智能、计算机科学和心理学。他正在重新定义和革新哲学家的角色。

迈克尔·舍默　美国头号“科学打假人”，怀疑论者学会创始人

丹尼特是我们这个时代最具原创性的思想家之一。



扫码获取“湛庐阅读”App，
搜索“直觉泵”，
获取丹尼特的演讲及其他精彩内容。



什么是彩蛋

彩蛋是湛庐图书策划人为你准备的更多惊喜，一般包括①测试题及答案 ② 参考文献及注释 ③ 延伸阅读、相关视频等，记得“扫一扫”领取。

推荐序一　从流俗的看法中跳脱出来

陈嘉映
著名哲学家，首都师范大学哲学系教授

《直觉泵和其他思考工具》一书推荐了一批“思考工具”。这些直觉泵，有些是作者自己的发明创造，有些取自前人。据作者所说，很多强有力的思考工具是数学性的，不过，这本书推荐的则是那些非形式化的思考工具，散文式甚至诗歌式的思考工具，这一点本来从“直觉泵”这个书名就能想到。

但这本书远远不止于推荐思考工具，实际上，在非形式化的王国里，在散文和诗歌里，原不存在单纯的工具。问题的解决方式不可能完全从相关问题那里抽象出来。我们时代直面的很多根本哲学问题，例如意识问题、自由意志问题，皆是如此。这本书差不多始终是依托这些具体问题来谈论方法的。他推荐各种思考工具，根本意图是帮助我们从流俗的看法中跳脱出来，其实，所谓“哲学”，干的就是这个。

跳出来，跳到哪里？丹尼特从来不忌惮最艰深的问题，而且对这些问题都有自己的解决方案，或至少，有他解决这些问题的思路。我个人未见得全部认同他的方案和思路，不过，他有大量的奇思妙想，能不断给人启发。也许他错了，但独到的错误、有意思的错误，不会像人所周知的正确那样让人扫兴。就算你不是个哲学爱好者，你也可能喜欢这本书，书里有很多有趣的知识，例如“生命游戏”，讲解得流畅易读。读过之后，你会感觉自己顿时长知识了。

我很多年前开始读丹尼特的作品，立刻感到这位作者有过人的求知欲，或更准确地说，有过人的求解欲。不管是你想得到想不到的问题，都会引发他的兴趣。顺理成章，他不餍足地学习科学，从其中搜索解答哲学问题的线索。当代哲学家里像他那样熟悉最新科学成果的人不多。但他并不盲信科学，认为科学家也会犯一些初级错误。丹尼特相信，有些问题是哲学问题，不是科学能够解决的，科学家一旦把它们混同于科学问题就难免想错，而且，由于他们的科学训练要求他们抛开常识，他们很可能比普通人错得更离谱。

丹尼特知道得太多、想得太多、说得太多，难免有点儿芜杂。我想

不出谁能一口气读好几本丹尼特的书，隔一段时间读一本似乎效果更佳，甚至不妨隔一段时间读他一章两章，让我们很容易消沉的脑神经重新活跃起来。

推荐序二　哲学家的法宝

万维钢
科学作家，“得到”APP《精英日课》专栏作者

这是一本能让你变得更聪明的书。丹尼尔·丹尼特的《直觉泵》现在已经是名著了，如果你拥有超一流的智慧，你想必已经知道这本书。如果你没读过，而你又愿意在脑力上下功夫，我建议你把这本书啃下来。

读完这本书，你的思维水平会有一个不可逆的提高。再听到别人一本正经的愚蠢言论，你会受不了。想起以前自己说过的那些随意的话，你会感到脸红。

要想横扫街头混混，你最好的办法是学习专业的搏击术。要想在思想上变聪明，你需要高级的思维训练——

学哲学。

哲学死了吗？

哲学，曾经被视为最高级的学问。很多中国人曾经相信，要想把什么事情做好，比如打仗、生产，甚至搞科学研究，你必须以某种哲学思想为指导。哲学高高在上，其他一切学问都只不过是哲学在各个领域的具体应用罢了。

但现在可没人敢这么说。至少在很多物理学家这里，哲学是被嘲笑的对象。以前人们不知道宇宙是怎么回事，某些哲学家空口无凭地就敢说“整个物质世界在时间上无始无终，在空间上无边无际”——在现在的物理学家看来这个世界观简直就是一种原始宗教观念。最新的物理学认为我们这个宇宙在时间上一定有个开端，在空间上至少曾经只有有限大。

鉴于哲学已经不再是世界观答案的合法来源，霍金在《大设计》这本书中写道：“哲学已死。”你想知道这个世界到底是怎么回事儿，你得问科学家。科学的答案不是一个科学家坐在那里空想出来的，也不是几个科学家辩论出来的，而是花费无数时间和金钱，通过无数个精巧的科

学实验得出来的。

费曼——他是公认最聪明的物理学家之一——对哲学有众所周知的敌意。费曼曾经说，什么是哲学家呢？就是这么两个人，一个哲学家对另一个哲学家说：“你根本不知道我说的意思！”另一个哲学家说：“什么是‘你’？什么是‘我’？什么是‘知道’？”费曼说我们能不能进行一点有建设性的讨论。[(1)]

杨振宁也鄙视哲学。别人问杨振宁哲学能不能指导科学，他说每个科学家都有自己的研究风格，如果你把这种个人风格称为“哲学”，那的确能；但是如果你说的是作为一门正式学问的哲学，那完全不能。有人又问，可是日本物理学家坂田昌一就很推崇哲学啊？杨振宁说坂田昌一要是“越少用哲学，他的成就越大”。[(2)]

坂田昌一的哲学思想包括“物质无限可分”。这个思想非常符合直觉，但是在现代粒子物理学家看来，所谓“无限可分”根本毫无意义。原子可以分成原子核和电子，原子核分成质子和中子，质子、中子和电子分成夸克，那夸克应该怎么分呢？事实上，物理学中的基本粒子，比如“夸克”，本质上是一个数学结构，根本就没有再“分”的必要。哲学家的原始直觉没触及这个情况。

所以哲学真的不能指导科学。真实情况是科学在指导哲学，是哲学家们从新科学发现中获得坐而论道的素材。

那哲学到底还有啥用呢？

大问题

我曾经是一个普通的物理学家。我不可能比霍金、费曼和杨振宁更高明，但是因为我是一个特别谦虚好学的人，所以我对现代哲学略有一点了解。

哲学没死。哲学不能指导科学，但是科学也不能取代哲学。有些问题，只能交给哲学家。

有些问题是在科学之外的。比如说，人体的工作原理，这是科学问题。那人怎样才能获得幸福？人到底应该怎么做才是道德的？人和人之间能平等吗？这些就只能是哲学问题。康德、洛克、休谟，以及中国的孔子、孟子，他们的思想并没有因为科学进步而过时。

有些问题是不需要使用科学手段就能解决的。比如说有个科幻电影叫《盗梦空间》，说可以通过梦境来修改一个人的观念。请问这是可能的吗？我们能给一个人的大脑输入任何信息吗？

答案是不能，而你不需要脑科学的知识就能想明白这一点。丹尼特这本书里有个直觉泵叫“生活在克利夫兰的一位兄长”，说的就是这个意思。如果这个人并没有兄长，你生生地在他头脑中植入一条“我有个生活在克利夫兰的兄长”的信息是没用的，因为他不知道这位兄长长什么样子、做过什么事、跟自己有什么关系，他甚至可能根本就没听说过克利夫兰这个地方：一种观念只能生长在一堆观念之上。

你不会在生理学或者心理学的教科书里找到这样的推理。这不是科学知识，这是哲学思辨。

只有哲学家能定义什么叫哲学——不过我估计哲学家对这个定义会争论不休——依我之见，所谓哲学，就是纯粹靠思辨来理解世界的学问。

有些问题，是科学想要回答，可是暂时回答不了的，而哲学思辨却能让我们有所探索。

比如说，如果组成人的原子因一直受物理定律左右而没有自由意志，那人能有自由意志吗？大脑的意识是怎么回事？计算机能有意识吗？意识这么精巧的东西，是怎么通过自然演化产生的呢？

这些是最高级的问题，无数科学家也在寻找答案，而哲学家可以帮忙。比如说，人刚出生的时候，大脑是一块没有任何信息的“白板”吗？丹尼特在书中使用一个叫作“身陷机器人控制室”的思想实验，说明人脑不可能像刚出厂的计算机硬件里一样一片空白，其中必定要预装一些软件系统。

事实上，现代职业哲学家对计算机科学、生物学和最先进的物理学的了解程度可能会让科学家震惊。他们完全知道科学进展到了哪一步，他们比科学家更善于把握全局，他们能把这些知识整合在一起，帮助我们理解这个世界。科学家的任务更多的是发现底层的原理和证实技术细节，哲学家则常常提供洞见。

学习哲学，你得有点钻研大问题的气魄。你得对自己的大脑有信心，相信脑子是可以用的。

哲学的技艺

科学家要做习题、棋手要打谱、律师和商人要学习案例，丹尼特说的“直觉泵”，就是哲学家的思维工具。这里的“直觉”是个好词儿，代表快速产生思维的灵感。直觉泵，就是说书中这些好像寓言故事和成语典故一样的思维工具，能给你“泵”出各种直觉来。

咱们举个简单的例子，看看有章法的思维和普通人的思维有什么区别。

现在比特币非常火了，要好几万元人民币一个，而几年前才值几块钱一个。有个技术爱好者很早就听说过比特币，当时没当回事儿。后来比特币几百元一个，他觉得太贵了还是没出手。现在的他追悔莫及，逢人就说他本来有机会成为百万富翁。

如果他学过一点哲学，他就不会说这种傻话。丹尼特讲了一个直觉泵叫“古怪的狱卒”。说有个监狱里的狱卒，会在每天半夜犯人们都熟睡的时候，把监狱的大门打开几个小时。那请问，在这几个小时内，犯人们是自由的吗？

如果那时候犯人能醒过来，他们的确可以逃跑！这就好像如果当年你知道比特币将来会那么值钱你一定会买一样。但是事实上那时候犯人是熟睡的，而你当年也真的不知道比特币的未来行情。

这个直觉泵的要点在于，当你谈论“自由”的时候，你必须考虑当时外界的因果背景。你不知道和你睡着了一样，你的因果背景里没有那个选项，那不是真的自由，那不叫“有机会”。

苏格拉底有句名言叫“未经审视的人生不值得过”，我们同样可以说，未经审视的思想不值得拥有——或者至少不值得说出来。

学习哲学的第一步，就是要学会审视自己的思想。

咱们再举个例子。一般人头脑中有个思维定式，就是把东西分类，特别是把人分成“我们”和“他们”。中国人如何如何，外国人如何如何，就算不是歧视，也总要有所区别。可是你是否审视过，人或者东西，真的有明确分类吗？

这本书中有个直觉泵叫“世界上第一只哺乳动物是什么”。什么叫哺乳动物，科学家有严格的定义。但即便如此，如果你考虑到生命演化，你会意识到哺乳动物和哺乳动物的原始祖先之间有一个漫长而模糊的中

间地带。你根本就没有办法来个“一刀切”，说这边就是哺乳动物、那边就不是。你会意识到世界上的很多东西并没有泾渭分明的分类界限。

那么下次要说“凡是……”的时候，你也许会更谨慎一点。

丹尼特是当今最著名的几位哲学家之一，他研究的是意识、生命、计算和智能这些大问题。这本《直觉泵和其他思考工具》，是一线哲学家送给外行和初学者的礼物。书中有很多工具是专门为了破解大问题而生的，相当于军用武器。不过就算这些武器不能都用在你的日常思维上，它们也能带给你极大的乐趣。

我认为哲学既不是规矩也不是火炬，哲学是技艺。如果你想用哲学的“结论”去指导科学或者生活，你很可能什么都得不到。但是如果你把哲学本身当作一门有意思的学问，并且在追逐大问题的过程中掌握哲学家的“思辨技巧”，你会得到法宝。

推荐序三　这是真正的批判性思维训练营

叶峰
著名哲学家，首都师范大学哲学系教授

这本书的作者丹尼尔·丹尼特是美国著名哲学家，长期担任美国塔夫茨大学哲学系讲席教授和认知研究中心主任，同时也是著名的公共知识分子，在学术圈子以外同样有很高的知名度。丹尼特一生致力于在当代认知科学和进化论的基础上解释人类心智、意识、自由意志及伦理道德的本质，同时致力于普及进化论、宣传无神论。一些媒体将他与理查德·道金斯（Richard Dawkins）、萨姆·哈里斯（Sam Harris）及克里斯托弗·希钦斯（Christopher Hitchens）等几位科学家、学者及作家并称为“新无神论四骑士”。

丹尼特著述丰富，仅从2003年以来就有5部独立完成及4部与他人合著的著作问世。他的一些较为通俗的作品机智风趣、深入浅出又不失严谨，在哲学学术圈子之外有很大的读者群，是适合非哲学专业人士阅读的高级科普及文化思想方面的畅销读物。他的专业性哲学研究在英美分析哲学界内部也有很大的影响，他是分析哲学界内公认的最出色的哲学家之一。

这本书讲什么

这本初次出版于2013年的《直觉泵和其他思考工具》，一方面列举了许多可用于探究心智、意识、自由意志等方面问题的思考工具，即所谓的“直觉泵”；另一方面也借助于阐明这些直觉泵概括了丹尼特本人一生中最主要的哲学观点。在前一方面，这本书试图激发读者对心智、意识、自由意志等方面问题的思考，可算是一本哲学启蒙读物，而在后一方面，这本书其实是丹尼特对自己哲学思想的一个总结，具有一定程度的专业性。

所谓“直觉泵”指的是一些思想实验，这些思想实验可以激活我们对某个事物或问题的一种直觉的认识，同时还可以对思想实验中的场景设置设想种种变化（即所谓的“转动直觉泵的旋钮”），以此来检验我们的

直觉是否合理。丹尼特指出，有些直觉泵激发了真知识，是好的直觉泵，但也有些直觉泵激发的其实是幻象，而我们可以通过“转动直觉泵的旋钮”来认识到那些是幻象。丹尼特在这本书的第5页就举了一个特别好、特别机智的直觉泵的例子，是伽利略提出的关于自由落体运动的思想实验。这里就不重复介绍了，因为读者只要愿意，读几页书就能读到它，欣赏到它的机智。

不过这里需要指出，在这本书中丹尼特主要关注的是一些与心智、意识、自由意志等有关的直觉泵，也包括一些与进化论有关的直觉泵，因为在丹尼特看来，理解进化是理解心智、意识、自由意志等的关键。所以读者不要将这本书误解为一本市面上常见的、一般性地介绍逻辑与批判性思维工具的教科书。它的书名有可能让人产生这种误解。虽然书名中没有提到进化论、心智、意识、自由意志等，但这本书的内容其实是围绕这些方面的问题展开的。书中详细讨论的直觉泵的例子都是有关这些方面的问题的例子，而没有详细讨论任何物理学、化学、经济学、社会学、历史学或政治学等方面的直觉泵的例子。所以在图书分类中，这本书也许更应当被归类于心灵哲学（philosophy of mind）类图书，而不是所谓的批判性思维（critical thinking）类图书。丹尼特写这本书的主要目的似乎是要阐释他自己的哲学思想，而不是教给读者一些一般性的思维工具。如果读者完全缺乏对进化论、心智、意识、自由意志等方面哲学问题的兴趣，而将这本书当作一般性的思维训练书籍来读，那么可能会对其中的内容不耐烦。

但也需要指出，那种泛泛地介绍批判性思维工具的教科书，其实很难深入地教会读者使用真正良好的批判性思维工具。真正学习掌握良好的思维方法不是靠读几个肤浅的例子就能做到的。它需要对一些作为样本的问题做很深入的分析研究，仔细钻研专家们对那些问题提出的种种相关的论证和反驳，包括一些相当复杂、相当微妙的论证和反驳，深入分析那些论证和反驳中包含的各种潜在的逻辑谬误，并尝试自己提出新的论证或反驳。只有在这样的很深入、相当专业、长时间持续的分析思考中，才能真正培养良好的思维习惯，掌握分析性思维、批判性思维的方法。

笔者本人有过数年的逻辑与批判性思维教学经验。在大学里教逻辑与批判性思维这门课，一般都是采用市面上常见的逻辑与批判性思维教科书。这些教科书中有大量例子说明有哪些论证手段，有哪些常见的思维谬误，等等。但那些例子中的论证一般都太浅、太简单，几句话就说

完，完全缺乏论证上的深度。由于读者或上课的学生的专业方向很不一样，那些逻辑与批判性思维教科书一般不会采用某个特定专业方向中的很深入、复杂的论证作为例子。这是无可奈何的选择。结果是，一门课学下来，学生在课后做练习题时，的确会识别一些简单例子中的思维谬误，会完成一些简单的论证，但他们自己在日常生活及专业学习中的思维习惯似乎并未受到任何影响，还是按老习惯去思考，原来易犯的思维谬误还在继续犯着。

那些逻辑与批判性思维教科书中用几行文字就能说完的简单例子，实在不能让读者或学生真正体会到什么是良好的分析性、批判性思维，不能让他们改进自己的思维习惯，培养正确思维的习惯。所以，只要读者对心智、意识、自由意志、进化论等方面的问题有兴趣，愿意很深入地思考这些方面的问题，相信这本书可以当作一本更好的培养分析性思维、批判性思维的书籍。

当然，这本书主要还是阐述丹尼特自己的哲学思想。丹尼特的哲学思想的背景，一方面是当代分析哲学学术研究中的自然主义思潮，以及在“自然主义”这个大框架下发生的有关意识、意向性、自由意志等方面的问题的争论，另一方面是美国社会文化中的无神论与基督教信仰之间的争论。下面就这两方面的背景做简单的介绍，希望能有助于读者理解这本书的内容。

自然主义思潮

很多不同的学科领域都用“自然主义”这个标签，这里说的自然主义是指当前国际分析哲学界的一种主流哲学立场，虽然“自然主义”这个词的含义即使在分析哲学文献中也不完全确定。支持自然主义的分析哲学家一般非常尊重科学，完全接受科学的结论，在科学结论的基础上进行自己的哲学思考。特别地，他们接受进化论，承认人类是自然进化产生的自然生物体。同时，他们不相信有上帝、灵魂、先验自我、绝对精神等等，他们认为这些都是超自然的实体。他们也拒绝传统的，笛卡尔式或胡塞尔式的第一哲学思辨，认为这些都预设了超自然的先验立场。根据2009年的一个问卷调查，英美国家哲学系中大约50%的教师明确表示自己支持或倾向于支持自然主义，而只有大约26%的教师明确表示反对或倾向于反对自然主义。所以自然主义是国际分析哲学界的一个主流倾向。

自然主义作为当前国际分析哲学界的主流倾向起源于20世纪初的逻辑经验主义，但它最初、最重要的鼓吹者是美国哲学家蒯因（Willard Van Orman Quine）。蒯因也许是对当前分析哲学研究现状影响最大的哲学家，他保留迄今的思想遗产及影响可能超过了弗雷格、罗素、维特根斯坦、卡尔纳普等这些早期的分析哲学开创者。蒯因是丹尼特在哈佛大学读本科的时候的老师，而且曾特别尽力地将丹尼特推荐到牛津大学，在另一位具有强烈的自然主义（及行为主义）倾向的哲学家吉尔伯特·赖尔（Gilbert Ryle）的指导下攻读博士学位。

丹尼特显然完全接受了自然主义的基本立场。不仅如此，与其他也接受自然主义立场的多数哲学家相比，丹尼特还更多地尝试引入认知科学和进化论的成果来研究哲学问题，更积极地将哲学科学化。早在1965年完成的他的博士论文中，他就尝试引入神经科学的成果来讨论心灵哲学问题，而且他一直将这种尝试坚持至今。注意，在20世纪60年代，认知科学与脑神经科学研究还处于非常原始的阶段，远远不像最近20年那么热门，那样受到哲学家们的关注。

但是在“自然主义”这面大旗下针对一些具体的哲学问题还是有不同的观点及争议的。丹尼特最关心并深入研究过的是其中的意向性、意识和自由意志问题。他的专业哲学研究主要是在这些问题上与他的同样具有自然主义倾向的同行们进行争论，论证他自己的独特的立场观点。至于他作为公共知识分子所创作的一些较具普及性的作品，其主要目的则更多地是站在自然主义立场反驳反自然主义的有神论，为无神论作辩护。这本《直觉泵和其他思考工具》则兼有普及性作品和专业研究论著的特征。下面将简要介绍一下这本书所探讨的意向性、意识和自由意志等方面问题的背景。

意向性问题

“意向性”这个哲学术语在哲学文献中有多种含义。在分析的心灵哲学中，意向性主要指的是我们的概念、判断似乎能够在某种意义上表征或指向世界中的事物及事态这个事实。这种表征能力就被称为意向性。比如，我们的“兔子”这个概念表征（或指向）兔子那一类事物，“兔子吃草”这个判断则表征兔子吃草这个事态。

心灵哲学中对此的争议，首先是关于人的大脑中究竟有没有“兔子”这个概念或“兔子吃草”这个判断。常识中我们似乎都认为有。但是，

考虑到大脑其实是一个非常复杂的神经元网络，而神经元网络对信息的记录很可能是所谓分布式的、整体性的，也就是说，有可能数以百亿计的神经元以某种方式互相联结在一起就整体上同时记录了“兔子吃草”“草是绿的”“邮筒是绿的”“学校门口有个邮筒”“明天开学了”“我不开心”等等所有这些无数的信息，而不能指出其中某个或几个神经元说，那就是“兔子”这个概念，或那就是“兔子吃草”这个判断。所以，有的哲学家认为，大脑中其实并没有“兔子”这个概念或“兔子吃草”这个判断，这些所谓概念、判断都只是我们的幻想。大脑中有的是超大一堆神经元整体地记录了关于世界的所有认识。这是关于意向性的一种听起来较极端的观点，被称为关于意向性的取消主义观点。它不是否认大脑能认识世界，大脑当然能。它是否认大脑中有“兔子”“老虎”等一个个的概念。持有这种观点的著名的代表人物有保罗·丘奇兰德（Paul Churchland）。

另一种极端的观点是，大脑中的确有“兔子”这个概念及“兔子吃草”这个判断。他们认为，大脑中有一种思想语言，思想语言中的一个语词可以是由一个或数个神经元构成的，而概念就是思想语言中的这种语词，判断就是由这些语词构成的语句，大脑中的思维过程就是对脑中的这些语词、语句进行变换、推理。他们认为，基于符号推理的人工智能系统真实或近乎真实地反映了大脑实际的工作方式。支持这一类观点的著名代表人物有杰里·福多尔（Jerry Fodor）。他们认为，大脑中的神经元以某种方式实现了语词、语句这种东西，以及对语词、语句的变换推理，就像电脑硬盘中的0-1二进制序列以某种方式实现了语词、语句以及对它们的变换推理。

丹尼特本人则持有某种折中观点，他自己称之为“意向立场”观。它的大意是说，一方面，我们可能不确定大脑中是否真的有实实在在的“兔子”这个概念或“兔子吃草”这个判断这种东西，但另一方面，我们在描述一个人的智能行为的时候，将其描述为具有概念、判断、信念、欲望等等是合理的、成功的、有效的。所以丹尼特的观点既注意到了这个事实：大脑是一个极其复杂的神经元网络，在其中不一定能挑出“兔子”这个概念这个特别的东西。同时他又照顾到了另一个事实：我们常识中用“概念”“判断”“信念”“欲望”等术语来描述、解释人的行为似乎是很成功的。究竟丹尼特的“意向立场”观点是否比其他观点更合理，或者它只是一种勉强的折中观点，读者不妨自己研读这本书的第二、三、四部分，这几个部分集中解释了丹尼特的“意向立场”观。

意识问题

关于意识问题，当前分析的心灵哲学中争论的焦点是所谓的“感受质（qualia）”。意识包括许多方面，比如，使得清醒区别于沉睡的清醒意识，使得觉察到一件东西有别于无意识地看见它的觉知，使得记起某件事有别于让它沉睡在记忆中的记忆意识。绝大多数心灵哲学家都承认，意识的许多方面，包括这里提到的清醒、觉知、记忆，都原则上可以归结为神经元活动的一些特征。比如，清醒区别于沉睡就在于大脑的神经元活动的某种特征，觉察到什么区别于大脑无意识地接收并处理某项信息也是在于神经元活动的某种特征。

但有一些心灵哲学家提出，具有一种感受（或称感觉）是一类非常特别的意识状态，它们无法归结为神经元活动，而且感受具有感受质，而感受质上的差异无法归结为神经元活动特征上的差异。所谓“感受质”指的是我们的主观感受的特质（quality），比如，疼痛的感受质不同于痒的感受质，刺疼的感受质不同于灼痛的感受质。那些心灵哲学家提出，这种主观感受质上的差异绝对不可能从神经元的角度得到解释。比如，也许有一天我们可以非常完备地描述刺疼与灼痛的神经元活动模式之间的差异。但即使如此，我们也没有解释为什么刺疼感觉起来是那样，而不是灼痛的感觉那样，为什么两种感觉之间有那样的差异。

这里我们只能用“那样”这种词来指称那种特别的感觉，因为我们不能像描述事物的形状、颜色、质量、分子结构等那样去描述那种感觉。一方面，只有自己体验过那种感觉才能知道它是怎样的，另一方面，即使你将一个人感受到刺疼时的神经元活动都描述得非常完备了，也没有说出那种感觉究竟是怎样的。所以这些心灵哲学家认为，感受及感受质是本质上非物质的，本质上不同于所有当前的物理、化学、生物学、生理学、神经科学等科学分支所研究描述的东西。感受质是只有靠主观体验才能把握的东西，而当前科学所研究的都是能够被客观地描述的东西。当代持有这种观点的比较有名的哲学家包括索尔·克里普克（Saul Kripke）、大卫·查尔莫斯（David Chalmers）等。这种观点被称作属性二元论。

有必要将这种属性二元论与传统哲学中笛卡尔式的实体二元论区分开。在笛卡尔的时代，人们无法想象大脑这一块物质如何能够思考，人们普遍接受的观念是：物质有广延但不能思考，灵魂能思考但没有广延。自从计算机人工智能出现以后，人们的观念发生了很大的变化。一

个智能机器人能够理解你的命令、观察环境、收集信息、进行推理、计划行动、执行计划、完成你交待的任务。它似乎显然具有概念、判断，能进行推理、计划等思维活动。所以今天，机器能思考这一点似乎是没有疑问了。

不过，虽然机器人的观察、推理能力很强，虽然它很聪明，但它是否有感觉呢？比如，一个人形机器人可能可以非常准确地模仿人们在感受到疼痛时的行为，包括修复被烫伤的皮肤、呲牙咧嘴等，但它是否真的会觉得疼？似乎不会，因为它内部就是一行、一行的人工智能程序在运行，在控制肢体的种种动作。似乎它不会有疼痛这种感觉。所以今天的属性二元论者的观念是：物质能思考，但没有感受。今天关注及争论的焦点从思想能力转到了感受能力。感受，而非思想，被一些哲学家认为是本质上非物质的。

丹尼特本人在这方面的观点一般被认为是一种极端立场，而不再是一种折中立场。他基本上是说，根本没有普通人及属性二元论者所说的那种所谓的感受，认为有那种感受其实是一种幻想。人脑中有的就是一些物质性的东西，即一些神经元活动，再没有其他的，尽管我们可以在不同层次上描述那些物质性的东西。属性二元论者尝试提出了一些直觉泵来论证存在那种非物质性的感受。这本书的第六部分主要就是试图论证那些所谓的直觉泵是虚假的，它们激发了幻象而非真知识。在很多人看来，丹尼特的极端观点是很反直觉的。可能我们多数人直觉上都认为，疼痛那种主观感觉不仅仅就是一些神经元活动，而是另外的、主观的东西。你不妨也读读这本书的第六部分，看看丹尼特能否说服你。

自由意志问题

在自由意志问题上，丹尼特似乎又回到一种比较折中的立场。关于自由意志的争议在于这样一个问题。根据科学，一个人就是一个生物-物理系统，他的所有行为都是被自然律决定的。比如，一个抢劫犯的抢劫行为的根源是他的基因以及他过去接受的环境刺激塑造了他当前的脑神经元结构和活动模式，这些脑神经元结构和活动模式控制他的肢体在一定环境中必然地做出了那个抢劫行为。既然一切都是这样被脑神经元和自然律决定的，那么他似乎不是出于他的所谓自由意志而做出了抢劫行为。而既然不是出于自由意志的行为，似乎我们就不能说他对他的抢劫行为负有道德或法律责任。这种结论自然在社会上多数人看来是不可接

受的。有一些科学家的确宣称自由意志是个幻象，人的一切行为都是被自然律完全决定的，人没有任何自由意志。这很让人绝望。

丹尼特在自由意志问题上持有的观点是所谓的决定论与自由意志的兼容论。兼容论者一方面承认，一个人的行为完全是受自然律决定的，没有所谓的超自然的灵魂或心灵来做出所谓的自由选择，但另一方面他们又认为，我们还是可以说人有自由意志（因此有道德和法律责任）。当然，这个“自由意志”也许不再是或不完全是你原来所想象的那种自由意志，至少肯定不是被设想为有着超越自然律的自由选择能力的灵魂所具有的那种自由意志。但是，这些兼容论者认为，至少我们还可以说人有自由意志，而且可以说，因此他们对自己的行为负有责任。很多当代哲学家都支持这种兼容论观点，但这种观点也自然面临来自两边的批评。否认人有自由意志的一边，以及相信人有更真实的自由意志的另一边都会说，以这种方式认为自己有自由意志可能只是廉价的自我欺骗与自我安慰。

自由意志问题是个非常令人困惑的问题。一方面，我们都本能地认为自己有自由意志，自己可以自由地决定做这个或做那个，否则就等于承认自己是一具行尸走肉，是完全受其他东西的控制而行动的。但另一方面，似乎大脑就是一个复杂的神经元网络，受生物-物理定律的支配在活动着，就像一台电脑（包括基于人工神经元网络的电脑），看不出自由意志会从其中哪里冒出来。丹尼特在这本书的第七部分借助于一些直觉泵解释了他的兼容论。至于能否消除你的困惑，需要你试着去读读才能知道。

怎么读这本书

简单介绍了这本书的内容的背景后，可能还有必要提醒读者一点。对一本分析哲学的书，我们应该永远抱着分析的、质疑的态度去读。不是将读到的作者的话当作真理记下来，而是时时试图去反驳作者，去指出作者的论证中的毛病。所以，如果你读这本书的时候觉得丹尼特某个论述有问题，或认为他没说到点子上，或他说得不够清晰明确，希望你不要气馁或放弃。相反，要尝试找出更明确的理由反驳他。很多哲学家并不赞同丹尼特的观点，你反对丹尼特的理由很可能跟其他哲学家的观点其实是相通的。当然，那些与丹尼特的观点相反的观点也不是无懈可击的。哲学所探讨的问题极为复杂，对每一方的观点都有很多支持和反

对的理由。因此，要综合所有这些理由，找到对问题的一个最合理的回答，一般是非常困难的。但这也正是哲学的魅力所在。不喜欢思考、质疑的人可能也不会喜欢分析哲学。这是一本可以帮助思维训练的书籍。正是在分析、质疑中才能锻炼自己的分析性思维、批判性思维的能力。相信丹尼特本人会很欢迎你去反驳他，指出他的论证中的弱点。

翻译这本书的三位青年才俊都有多年的学习与研究分析哲学的经历。他们的本科专业背景有来自文学的，也有来自理工科的，对分析哲学的浓厚兴趣将他们聚到一起，在首都师范大学哲学系跟随陈嘉映和夏年喜教授读研究生。其中，冯文婧博士已经毕业，目前正在做博士后研究，她的博士论文正是专门研究丹尼特哲学思想的，她熟悉丹尼特的所有哲学著作，还有很流畅的文笔。傅金岳是陈嘉映教授的在读博士，但其实在读博士之前他已经学习与研究分析哲学很多年，非常熟悉当代心灵哲学的发展路径及相关论题的探讨，中英文水平俱佳，也有不少翻译经验。徐韬曾是陈嘉映教授的硕士研究生，自2013年毕业起，一直从事与西方哲学相关的写作与编纂工作。他们的译本很大程度上保留了丹尼特作品风趣通俗的风格，非哲学专业的读者应该也很容易理解这个译本。

在学术界内，翻译是一项吃力不讨好的工作，很耗费时间精力，又不能算真正的研究成果。但对这本书及同类的传播、普及思想与知识的外文书，可能只有通过翻译才能让国人真正受益。虽然目前英语很普及了，但国内的普通读者，即使英文阅读能力不错，直接读英文也相对更吃力些，需要花更多的时间。在这个工作、生活节奏异常紧张的年代，恐怕很少有国人能够去阅读那么厚的一本与自己的工作不那么直接相关的英文书，很少有国人能够仅仅为了长点见识、提升思想，就抱起一本近五百页的大部头的英文书，认真地读下来。因此，翻译真的对很多人有很大的帮助。如果没有译本，很多人可能就下不了决心花那么多的时间去读那些书。笔者受三位译者之托代写这篇序言，非常希望他们的努力，他们花费时间、精力的成果有人欣赏。

献给塔夫茨大学，我的学术家园

前言　我要写一本人人都能读懂的书

塔夫茨大学是我40多年以来的学术家园，对我来说，它就像童话中金发姑娘那碗刚好可口的粥一样，似乎总是恰到好处：没有太多压力，也不让我太过放纵；让我拥有可以轻松讨教的同僚，他们才华横溢却毫无恃才傲物之心；让我有足够认真的好学生，他们值得厚待却不需要全天候去培养；它让我明明身在象牙塔却可以去自由地解决现实世界中的问题。塔夫茨大学自1986年创建认知研究中心以来就始终支持我的研究，让我从求资筹款的苦役中解脱出来，从而有充分自由在诸多领域实现与众人的合作：或者是远游参加工作坊、实验室和会议，或者是将学者和专家们邀请至研究中心访问。这本书展示了我这些年一直在做的事情。

2012年春天，我在塔夫茨大学哲学系举办的一次研讨班上试讲了本书各章节的初稿。试讲是我多年来的习惯，但这一次，我想借助学生们的帮忙，让这本书尽可能地被大众读者所接受，所以，我把研究生和哲学专业的学生排除在外，只选前12个报名的无畏新生组成讨论班。后来由于办公人员的失误，多选了一个，所以实际上是13个。我们彼此引导，非常愉快地展开书中相关话题的讨论，他们知道，他们真的可以与教授顽强对抗，而我更是确定，我真的可以回溯到更基础的层面，更好地解释这一切。感谢这些年轻合作者们的勇气、想象力、能量和热情，他们是：Tom Addison、Nick Boswell、Tony Cannistra、Brendan Fleig-Goldstein、Claire Hirschberg、Caleb Malchik、Carter Palmer、Amar Patel、Kumar Ramanathan、Ariel Rascoe、Nikolai Renedo、Mikko Silliman和Eric Tondreau。

经研讨班探讨所得的第二稿，之后又由我的几位好友博·达尔布姆（Bo Dahlbom）、休·斯塔福德（Sue Stafford）和戴尔·彼得森（Dale Peterson）通读，他们进一步为我坦诚地提供了有用的评估和建议，其中大部分我都接受；同样通读过第二稿的还有我的编辑德雷克·麦克菲力（Drake McFeely），他在W. W.诺顿公司布伦丹·柯里（Brendan Curry）的大力协助下，还负责了这本书的许多改进工作，对此我深表感激。特别鸣谢认知研究中心的项目协调员特雷莎·萨尔瓦托（Teresa Salvato），她不仅以各种各样的方式为这整个项目提供了直接帮助，还

因为非常有效地料理着研究中心和我的行程，让我有更多的时间和精力制作和试用我的思考工具，间接推动了项目进展。

最后，一如既往，我要表达对我妻子苏珊（Susan）的感谢和爱意。我们并肩作战已有50年，对于我们一起合作的事情，她和我一样认真负责。

目 录

推荐序一　从流俗的看法中跳脱出来
推荐序二　哲学家的法宝
推荐序三　这是真正的批判性思维训练营
前言　我要写一本人人都能读懂的书
引言　什么是直觉泵
第一部分　通用思考工具
01　犯错儿　犯“好”错儿才有价值
02　归谬法　发现错误命题的妙招
03　拉波波特法则　批评他人的正确方式
04　史特金定律　不要把时间浪费在无意义的事情上
05　奥卡姆剃刀　如无必要，勿增实体
06　奥卡姆扫把　有意隐瞒对自己不利的证据
07　外行做媒　既不“过少解释”，又不“惹恼行家”
08　跳出系统　打破惯性思维的好方法
09　古尔德的3种思考工具　“不如说”“故意堆积”和“古尔德两步”
10　小心“当然”这个词　一种让你无需思考就认同的花招
11　反问　让你不好意思说“不”
12　什么是“深马”　爱就一个字
第一部分小结
第二部分　关于意义和内容的思考工具
13　特拉法尔加广场上的谋杀案　意义和内容都是我们大脑的“意向性”产物
14　生活在克利夫兰的一位兄长　一种观念只能生长在一堆观念之上
15　“爸爸是名医生”　理解是随着时间推移而逐渐清晰的
16　常识映像和科学映像　反映世界的两个不同视角
17　常识心理　人们无需正规教育就拥有的一种能力
18　意向立场　解释实体行为的一种策略
19　人与“次人”的区别　自上而下的认知能力分解
20　大脑中的小人儿委员会　认知科学的好理念
21　近似算子　达尔文渐进主义的启示
22　神奇组织　不要为逃避难题而求助于“魔法”
23　身陷机器人控制室　初长成的大脑不是一块“白板”
第三部分　关于计算机的思考工具
24　计算机施展魔法的7个秘密　部分之和大于总体的最好实例
25　虚拟机　模仿硬件运行的计算机程序

26　算法　有效解决问题的策略机制
27　让电梯实现自动控制　机器取代人类的逻辑
第三部分小结
第四部分　更多关于意义的思考工具
28　红发人那事儿　大脑中的认知机制
29　彷徨的双币机、孪生地球以及巨型机器人　原初意向性和派生意向性存在明显的边界吗
30　彻底翻译与蒯因式填字游戏　不存在绝对正确的翻译
31　语义引擎和句法引擎　大脑只是通过句法引擎模仿语义引擎
32　沼泽人遇上母牛鲨　哲学家最喜爱的直觉泵
33　两个黑盒子　究竟是什么让红灯闪烁
第四部分小结
第五部分　关于进化论的思考工具
34　万能酸　摧毁一切的达尔文思想
35　孟德尔图书馆　只有极少数的A、C、G和T组合才是有意义的
36　什么是基因　基因就像软件中的子程序
37　生命之树　我们人类与谁有着共同的祖先
38　自然选择VS智能设计　“起重机”VS“天钩”
39　小布谷鸟为什么会把宿主的蛋推出鸟窝　无需理解的能力之一
40　白蚁城堡是谁设计的　无需理解的能力之二
41　蝉知道素数的含义吗　无需理解的能力之三
42　瞪羚的弹跃　无需理解的能力之四
43　世界上第一只哺乳动物是什么　压根儿就不存在
44　物种形成于何时　某个精确但又不为人知的时刻
45　回溯性加冕　人类最近的母系祖先“线粒体夏娃”大约生活在20万年前
46　循环　周期性重复是进化的关键
47　青蛙的眼睛究竟将什么告诉了青蛙的大脑　扩展适应的最好例子
48　穿越巴别图书馆　福尔摩斯的高效推理
49　谁是《帕姆雷特》的作者　“延伸的表现型”为你揭开谜底
50　虚拟旅馆的噪音　妙用“意外”是创新的关键
51　哈尔是赫布和爱丽丝的孩子吗　克雷格·文特尔说：是
52　模因　不论好与坏，它都是我们人类的共生体
第五部分小结
第六部分　关于意识的思考工具
53　对立的两幅意象　一幅错误的意象要用另一幅相反的意象来抵消
54　僵尸直觉　莱布尼兹的不合理推论
55　僵尸和殭尸　如何构想出一个哲学僵尸
56　花椰菜的诅咒　感受质的意义之一

57　“美元活力”究竟值多少钱　感受质的意义之二
58　卡普格拉先生的悲惨境况　感受质的意义之三
59　听声辨牌　感受质的意义之四
60　“中文屋”错在哪里　理解可以从一堆能力中冒出来
61　从火星来的远程克隆人　他还是原来的那个“他”吗
62　叙事重心　究竟什么是自我
63　他者现象学　人类语言的妙用
64　色彩科学家玛丽：揭露吊杆托架　错误假设导出错误结论
第六部分小结
第七部分　关于自由意志的思考工具
65　一位恶毒的神经外科医生　非手术的手术
66　生命游戏　我们的世界是被决定的吗
67　石头、剪子、布　最好的套路就是没有套路
68　两种彩票　从新视角看“决定论”
69　无效历史事实　用强行分类抹平微小差异
70　一场计算机国际象棋马拉松　如果决定论是真的，还存在真正的选择吗
71　终极责任　斯特劳森的错误论证
72　掘土蜂状　有限能力无法成就自由选择
73　巴西来的男孩　一个干扰人们思考的“坏”直觉泵
第七部分小结
第八部分　做个哲学家是一种怎样的体验
74　浮士德式交易　哲学家的目标不同于科学家的目标
75　简单的自我人类学　哲学家应该把反直觉的一切都去掉
76　“像棋”中的二阶真理　若不值得做，亦不值得做好
77　只需关注那10%的精品　哲学领域的史特金定律
结语　加倍努力地使用这些工具吧
附录一　未写入本书正文中的几个直觉泵
附录二　寄存器机练习题的答案
素材来源
参考文献
译者后记

引言　什么是直觉泵

你不能空着手做木工活，更不能空着脑袋思考。

——博·达尔布姆

思考是件难事。有些问题艰难得让你一思考就头疼。我的同事、神经心理学家马塞尔·金斯伯恩（Marcel Kinsbourne）说，思考之所以困难，是因为通向真理的崎岖小径与诱人的捷径争持不下，而捷径往往只是条死胡同。我们在思考时所做的大多数努力就是在抵御诱惑。我们不断受到它们的纠缠，还必须为了手头上的任务硬着头皮想下去。唉……

有一则关于约翰·冯·诺依曼的轶事。冯·诺依曼是一位数学家也是一位物理学家，他将阿兰·图灵的想法，也就是所谓的图灵机，变成了真正的电子计算机，现在我们称之为冯·诺依曼机，例如你的笔记本电脑或智能手机。冯·诺依曼是一位大师级的思想者，以能够在头脑中闪电般地进行大量计算而闻名。像所有著名的故事一样，关于他的故事当然也有许多版本。按照我这个版本的说法，有一天，一位同事拿了一道智力题给他，这道题有两种解法，一种是需要复杂计算的繁琐解法，另一种是比较巧妙的解法，是那种“啊哈”式的解法。这位同事有一个理论：遇到这道题时数学家会采用那个繁琐的解法，而更懒但更聪明的物理学家会停一下，然后找出巧妙的解法。那么，冯·诺依曼会采用哪种解法呢？

你们应该听说过这样一道难题：有两列火车相距100千米，在同一条轨道上相向行驶，一列火车的速度是每小时30千米，另一列的速度是每小时20千米。当两列火车相距100千米时，一只鸟以120千米的时速开始从火车A飞向火车B，到达后再飞回火车A，如此往复直至两列火车相撞。当两列火车相撞时，鸟一共飞了多远？“240千米。”冯·诺依曼脱口而出。“该死的，”他的同事说，“我猜你会用那个难的方法呢，求无穷级数。”“啊！”冯·诺依曼拍着自己的脑门尴尬地叫道，“原来还有一种简单的方法呀！”（提示：两列火车相遇所用的时间是多少？）

有些人是天生的天才，就像冯·诺依曼，他们可以轻松解开最棘手的谜团；而另一些人虽然步履蹒跚，却天赋“意志力”，这一英雄式的品质使他们能在追逐真理的道路上坚持到底。然后，就剩下我们这种人：没

太多天分还有点懒。可是我们仍渴望理解周遭的事物。该怎么办呢？我们可以使用大量的思考工具。这些方便的、辅助性的工具可以帮助我们拓展想象力、保持注意力，让我们妥当甚至优雅地思考真正的难题。

这本书就收集了我所喜爱的种种思考工具。我不只要描述它们，还将用它们引领你们的头脑轻轻穿过那令人不适的领地，直达关于意义、心灵和自由意志的本源之域。我们将从一些对各种问题都适用的简单而普遍的工具开始，其中一些我们已然谙熟了，而另一些尚未得到充分的关注和讨论。然后，我会让大家了解一些为特别目的而设计的工具，它们可以破除这样或那样误导人的想法，这些想法就像深深的车辙，连专家有时也会狼狈地深陷其中。我们也会认识并拆解种种不好的思考工具，因为一不小心，那些拙劣的说服技巧就会把我们引入歧途。无论你们是否能够安然抵达我所设定的终点，是否决定与我一起在那里驻留，这趟旅程都将用全新的思考方式武装你，让你得以思考这些主题、思考思想本身。

物理学家理查德·费曼（Richard Feynman）可能是比冯·诺依曼更富传奇色彩的天才，他的确天生就有一个世界级的大脑。他也爱做有趣的事，我们都该感激他，因为他特别乐于给我们透露那些使自己的生活变得更轻松的诀窍。不论你有多聪明，但凡有简单的方法可用，你就会变得更聪明了。他的自传《别闹了，费曼先生！》（*Surely You're Joking, Mr. Feynman!*）和《你干吗在乎别人怎么想？》（*What Do You Care What Other People Think?*）应该被每个有抱负的思考者列入他们的必读书目，因为这两本书给了我们很多启示，比如关于如何驾驭最棘手的问题，甚至还有在想不出什么更好的点子时如何花言巧语地把听众侃晕的方法。在受到他书中许多有益观点的启发，并且看到他在分享自己的大脑如何工作这件事上所表现出的真诚之后，我决定在同一个话题上试着谈谈自己的想法。它们没有多少自传性质，但却志在说服你们在思考那些论题时采用我的方法。我没什么诀窍，只是不厌其烦地诱导你们放弃对一些信念的固执，我的主要目标之一就是展示我想的是什么以及为什么这么想。

像所有工匠一样，铁匠也需要工具，但有句实际上已经没多少人知道的老话说得好："只有铁匠的工具自己造。"木匠不为自己制作锯子和锤子，裁缝不给自己制作剪子和针，水管工不给自己制作扳手，但是铁匠能以铁为原料制作锤子、钳子、铁砧和凿子给自己用。那么思考工具呢？谁制作了它们？它们是用什么制成的？哲学家制作过一些最好的思

考工具，不用任何材料而仅凭观念，即信息的有效结构。笛卡尔给我们带来了笛卡尔坐标系，如果没有坐标系中的x轴和y轴，那么由牛顿和莱布尼兹同时发明的最卓越的思考工具——微积分就几乎不可设想。帕斯卡为我们带来了概率论，这样我们才能简便地计算各种赌注的概率。天才的数学家托马斯·贝叶斯（Thomas Bayes）给我们带来了贝叶斯定理，它是贝叶斯统计思想的基石。不过，本书所着重讨论的大多是更简单的思考工具，不是数学和科学所用的精密、系统化的机器，而是头脑的工具。它们是：

标识。有时，在你翻来覆去思考某个事物时，给它取个生动的名字，会有助于理解这个事物。正如我们将会看到的，在所有标识中最有用的是警示性的标识或者“警报”，它们让我们对可能的错误来源保持警惕。

例示。有些哲学家认为，如果在哲学工作中使用例子的话，即便不完全是耍花招，至少也是不必要的。这颇像小说家在他们的小说中避免加入插图一样。小说家以用文字实现所有表达为荣，而哲学家则自豪于能够以严密的形式呈现仔细摆弄过的抽象概括，尽其所能地使其接近数学证明。他们干得不错，但可别指望我会把他们的作品推荐给我的学生，只有少数特别优秀的学生除外。他们的书原本不必这么难。

类比与隐喻。将一个复杂事物的特征与另一个你自以为熟知的复杂事物的特征相对照，是一种非常有力的思考工具。但由于它过于有力，一旦思考者们的想象被一个靠不住的类比所困，他们就会误入歧途。

脚手架。你可以只用一架梯子来盖屋顶、刷房子或者修理烟囱。你把梯子挪到这儿、爬上去，再挪到那儿、爬上去，每次只能在一个地方干活；但如果你事先搭起一些坚固的脚手架，你就可以在整个施工过程中灵活而安全地走动，最后干起活来反而更方便。本书所提到的一些特别有价值的思考工具就类似于脚手架，虽然把它们搭好需要费一些时间，但一旦搭好，就可以将很多问题一并解决，根本不用像挪梯子那样搬来搬去。

直觉泵。就是那种被我称为“直觉泵”的思想实验。

毫无疑问，思想实验是哲学家们钟爱的思考工具之一。如果我们可以通过一些巧妙的演绎推理得出一个问题的答案，谁还需要实验室呢？从伽利略到爱因斯坦还有后来的那些科学家们都用思想实验取得了良好的效果，所以思想实验并不为哲学家独享。有些思想实验像严格的论证一样是可分析的，常常具有归谬法（*reductio ad absurdum*）的形式，也就是一方往往利用对方的前提推出一个形式上的矛盾或者一个荒谬的结果，以此来说明对方的推论存在缺陷。

我最喜欢的一个思想实验来自伽利略，这个实验证明了在忽略摩擦的情况下，重的物体的下落速度并不比轻的物体快。他论证说，如果重的物体下落的速度比轻的物体快，那么重的石头A的下落速度就大于轻的石头B，如果我们把B系到A上，那么在下落时B就会像一个赘物那样拖慢A的速度。但是，A和B系在一起之后的总重量大于A，因此两者一同下落的速度也要大于A自身的下落速度。于是，我们得出结论：把B和A系在一起之后，它们的下落速度既大于又小于A的下落速度。这就出现了一对矛盾。

另一些思想实验没有这么严格，却往往同样有效：有些思想实验的小故事设计得令人感到心有灵犀、拍案叫绝，无论辩护的是哪个论题，人们看过都会觉得“是，当然，一定是这样！”我把这些思想实验称作“直觉泵”。我在第一次公开评论哲学家约翰·塞尔（John Searle）的思想实验“中文屋”（Searle, 1980; Dennett, 1980）时发明了这个术语，有些思想家认为我的这个术语有贬低或轻蔑的意味。但是恰恰相反，我爱直觉泵！也就是说，有些直觉泵是极好的，有些是有问题的，完全不可靠的只是少数。

几百年来，直觉泵在哲学中占据统治地位。它们是哲学家版的《伊索寓言》，早在哲学家出现之前，人们就已经把寓言视作一种极好的思考工具了。(3)如果你曾在大学里学过哲学，或许已经接触过这类经典之作，比如柏拉图在《理想国》当中所提到的“洞喻”：在洞穴中，人们被锁链拴住，只能看到真实事物投射在墙壁上的影子。再比如他在《美诺篇》（*Meno*）中所讲的苏格拉底教一个奴隶男孩学习几何的故事。还有“笛卡尔的恶魔”，在恶魔的欺骗下，笛卡尔相信了一个完全虚假的世界，这是最早的“虚拟现实”思想实验。以及霍布斯的“自然状态”，在这种自然状态下，人的生活是卑污、残忍和短寿的。虽然它们不如《伊索寓言》中的《狼来了》《蚂蚁和蚂蚱》那么著名，但也广为人知，其中的每一个都能调动起我们的某些直觉。柏拉图意在用“洞喻”启发我们思

考感觉与现实的本质，以奴隶男孩的故事为例来向我们说明先天的知识；恶魔是一台终极的“怀疑发生器”；而霍布斯关于“自然状态”的寓言旨在让人意识到，我们只有订立契约才能从自然状态中走出来。它们是哲学中永恒的旋律，余味深长，让学习哲学的人们在多年之后仍能相当生动而准确地记住它们，即便他们已经淡忘了那些复杂难解的论证和分析。

一个好的直觉泵比任何一种论证和分析都更为有力。我们将要思考种种当代的直觉泵，包括那些质量不合格者。我们的目标是理解它们好在哪儿、它们是如何工作的，以及我们如何使用甚至制作它们。

我可以以一个小故事《古怪的狱卒》为例。有个狱卒，每天夜里，他等到所有囚犯都熟睡以后，就会挨个打开所有牢房的门，一开就是几个小时。请问：这些犯人是自由的吗？他们有逃跑的机会吗？不见得。这是为什么？还可以以另外一个小故事《垃圾箱里的珠宝》为例。假设有一件价值连城的珠宝被人丢弃在了路边的垃圾箱里，而你在夜里散步时正好经过了那个垃圾箱。看上去你似乎得到了一个千载难逢的发大财的机会，但是它简直太“绝”了，绝得很难再算作是一个机会，你绝无可能发现这个机会进而采取行动，或者说你压根儿就不会往这方面想。

以上两个简单的情景所激发出的直觉是我们平时注意不到的：所谓一个真正的机会，重要的是我们可以及时得到有关它的信息，留给我们足够的时间注意到它并能容许我们为此做些什么。当我们渴望做出“自由”的、不由任何外力所导致的选择时，我们常会忘记，人不应该希望与一切外力相隔绝，因为自由意志与那个嵌入在丰富的因果背景中的我们并不矛盾，它实际上需要这个背景。

希望你已经发现了，关于自由意志我们还有很多东西可说！这些小小的直觉泵能把问题搞得活灵活现，不过目前它还什么都没有解决（后面有一个部分将集中讨论自由意志的问题）。我们必须慢慢地熟练掌握一些技巧，比如谨慎地使用思考工具，注意自己走到了哪一步，还要时时检查是否有陷阱。如果可以把直觉泵设想成是一种为了说服我们而精心设计的工具的话，那么，我们就应该能够通过逆向工程检查它所有的活动部件，看看它们是如何工作的。

1982年，当侯世达（Douglas Hofstadter）[4]和我合著《心我论》（*The Mind's I*）这本书时，他提出了一个恰到好处的建议：把直觉泵设想为一种包含多种设置的工具，我们可以“转动各个旋钮”，看看当设定发生变化时，相同的直觉还能否被它们激发出来。

下面我们就来鉴定一下，转动“古怪的狱卒”的旋钮吧。我们可以假设组成那个思想实验的每个短语都各有其功能，通过替换掉它们或者做出小小的改动，可以看出它们各自所发挥的功能。

1．每天夜里

2．他等到

3．所有囚犯

4．熟睡之后

5．挨个打开

6．所有牢房的门

7．一开就是几个小时

一种可能的改动是：

> 一天夜里，他命令守卫给一个囚犯下了麻药，当守卫们离开时又碰巧忘了关上那间牢房的门，一开就是几个小时。

这样改，情景的味道是不是变了很多？这是为什么呢？原版的要点仍然表达出来了，只是效果不如从前。最大的不同似乎是：把囚犯可能随时醒来的自然熟睡，改成被下了麻药。另一个改动是“碰巧”一词，这个词突出了有意与无意在狱卒与守卫的行为中所扮演的角色。而把“每天夜里”修改为“一天夜里”则似乎改变了概率，这对囚犯不利。你可能会问：什么时候又跟概率扯上关系了？关概率什么事？想一想，为了躲过一次“中奖者会被枪毙，而参加者上百万人”的乐透抽奖，你愿花多少钱？再想一想，为了躲过一次要用左轮手枪玩的俄罗斯轮盘赌，你又愿花多少钱？为了阐释一个直觉泵我们用到了另外一个直觉泵，这是个值得记住的小窍门。

另外一些旋钮则不那么明显：“恶毒的房东”等到房客们入睡后悄悄地锁上卧室的门。“医院主管”担心可能发生火灾，在夜里悄悄把所有房间和病房的门锁打开，但为了让病人们睡得安心，就没有告诉他们这件事。或者，假如我们把故事中的“监狱”变得比平常的监狱大些，有澳大利亚那么大，又会如何？你既不可能锁上也不可能打开澳大利亚所有房间的门。这会改变什么？

在我们处理任何直觉泵时，都应该带着这种自我意识上的谨慎，这种自我意识本身就是一种重要的思考工具。因此哲学家最喜爱的策略就是“溯元”（going　meta），即关于思考的思考、关于谈话的谈话、关于理由的理由。元语言是我们用来谈论其他语言的语言，元伦理学是居高临下检验其他伦理理论的理论。就像我跟侯世达说过的那样：“不论你研究什么，我都能搞出一个‘元什么’。”当然，这整本书就是“溯元”的一个实例：探索如何谨慎地思考那些帮助我们谨慎思考的方法，以及谨慎思考谨慎思考的方法的方法，等等。(5)2007年，侯世达给他偏爱的一些小工具列了一张清单：

追野鹅

黏合剂

鬼把戏

酸葡萄

力气活儿

泥足巨人

松动的大炮

狂想家

空头支票

大力扣篮

反馈

如果你已经熟知上面所说的这些，那就不会把它们只当作“一些词语”。上面写的每一条都指代一种认知工具，就像长除法和求平均数法一样。每种工具都能在一个较为宽泛的语境下发挥作用，有的工具能让一些假说更清楚地表达出来以方便人们检验，有的工具可以方便人们认识到一些不易察觉的规律，而有的能够帮助人们找到事物之间主要的相似点，如此等等。你的词汇表里的每一个词都是一件简单的思考工具，只不过有些工具比别的工具更有用。如果上面列出的那些工具，你的工具箱里一件都没有，难道你还不想要一些吗？有了这些装备，我们就能思考那些以前可能连表达都表达不清楚的思想。当然，就像古语所言：“如果手上只有锤子，就看什么都像钉子。”因此这些工具也都有被滥用的可

能。

让我们先看看其中一个工具是怎么被人滥用的吧：酸葡萄。它出自《狐狸与葡萄》的故事。这个故事提醒我们，人们总会假装不在乎那些自己得不到的东西。可以想一想，当一个人听到别人问她“你是吃不着葡萄说葡萄酸吗”时会有什么反应。她听完以后也许会发现自己真的是这么想的，这可能很有效地让她改变了自己的想法，或者让她站在一个新的角度来反思这件事。但是也有可能，这句话会明显让她感到自己被侮辱了。工具也能当作武器。《狐狸与葡萄》故事中的道德含义已经深入人心，你可能已经很难想起这个故事想要启发我们什么，也很难再感受到那些微妙之处了，虽然有时候它们并不重要。

得到一个工具与明智地使用它是两种截然不同的技巧，但是你得先从得到工具或者自己制造工具开始。我后面展示给大家的很多思考工具大都是我自己发明的，另一些是我从其他人那里得来的，我会及时向那些人致谢。(6)在侯世达刚才列的清单里，没有一件工具是他自己发明的，可他却为我的工具箱增添过不少好东西，比如跳出系统（jootsing）和掘土蜂状（sphexishness）。

很多最有力量的思考工具是数学化的，但在本书中我只会稍提一下，不费更多的笔墨。因为这本书旨在赞美那些非数学化的思考工具的力量，赞美那些非形式化的、散文和诗歌式的思考工具的力量，如果你愿意的话，也可以说，这本书所赞美的是那些通常被科学家们所低估的力量。你应该能猜到为什么。

首先，在研究性的期刊中有一种科学写作的文化，这种文化提倡作者用一种非个人化的方式简明扼要地陈述问题，尽量不用华丽的辞藻、修辞和典故。我们有充分的理由在最严肃的科学期刊上把文章写得冰冷乏味。就像1965年在牛津，我的博士考官、神经解剖学家J. Z.杨（J. Z. Young）在读了我的论文中那些在他看来有点古怪的散文部分之后所写到的：“英语现在已经成为国际科学界的通用语言，所以我们这些以英语为母语的人有必要把文章写到能让‘一位有耐心的中国人查阅词典’就能读明白的程度。”要知道我写的是哲学论文，并不是神经解剖学论文。这种自我强化的学术规范所导致的后果不言自明：不论你是中国、德国、巴西还是法国的科学家，你最重要的著作都必须用英文出版。你需要写一些干巴巴的、翻译起来一点困难都没有的英文，尽可能少用文化典故，避免意义微妙的词，不玩文字游戏，甚至最好连比喻都不用。在此种国际学术体制下所实现的一定程度上的相互理解当然非常宝贵，但是

也有代价：有些思想显然需要依赖日常比喻的传达，需要扭转读者的想象，需要利用书中的奇思妙想来冲破封闭我们心灵的藩篱。如果其中的某些言语不好翻译，那我也只能将希望寄托于技巧卓越的翻译家，或者希望世界各国的科学家把英语说得越来越流利了。

科学家之所以对“仅仅通过词语”进行的理论讨论抱有怀疑，还有另外一个原因，他们意识到：评判一个不用数学方程式来表述的论证往往要更难，更缺乏确定性。数学语言的说服力值得信赖，它就像篮筐上挂着的篮网。在操场上打过篮球的人都知道，如果篮圈上没有网的话，判断一个球是三不沾还是空心入筐会有多困难，而篮网从根源上消除了大家对“球进了还是没进”的争议。但有时候，有些问题太过狡猾，太过变幻莫测，无法用数学语言来表达。

我始终会考量，自己所做的研究能否向一群聪明的本科生解释清楚。我并不确定自己为什么要这样做，但这一挑战影响了我的所有作品。有的哲学教授特别愿意教那些只有研究生参加的高级讨论班，但我不是这样。研究生总是急于向同学们或者自己证明：他们已是行家里手了。他们熟练又自信地挥舞着本专业的行话来忽悠外行，炫耀自己能够小心地穿过那最曲折也最痛苦的技术化论证而不会被绕晕，因为只有这样才能让他们确信自己所做的事是需要专业技巧的。为高年级的研究生和各路专家所写的哲学书籍，一般来说全都没有可读性，所以它们大多数都没被人读过。

我这种试图把论证与解释描述得让哲学系以外的人也能够理解的写法有一个奇怪的副作用，它让有些哲学家在“原则上”不把我的论证当回事！很多年以前，我受邀去牛津大学的约翰·洛克讲座演讲，现场人满为患。听说有一位著名的哲学家在离席时嘟囔道：“要是我能从这种想在洛克讲座上吸引外行眼球的家伙身上学到东西，那我就完了。”果然，就我所知，他没有从我这儿学到任何东西。我不会改变我的风格，也永远不会为付出的这点代价而后悔。

在哲学里，有时你需要去做那种所有前提都带着编号，并遵循指定好的推理规则的严谨论证，但是这些其实没必要拿出来向公众展示。我们会要求自己的研究生在毕业论文里做论证，可是很不幸，有的人一旦养成这种习惯之后就再也改不过来了。不过我也要说句公道话，物极必反，像欧洲哲学那样天马行空的修辞，夹杂着文学性的藻饰和隐晦的暗示，对哲学来说也没有什么好处。如果让我必须选择一个的话，我宁愿每次都选难对付的善于分析的诡辩者，也不愿意听华而不实的夸夸其

谈。至少你能听懂前者讲的是什么，能够判断他说得对不对。

我相信，诗歌与数学之间难以界定的中间地带才是哲学家最能施展拳脚之处，他们会为真正艰深的问题提供清楚的解释。任何数学方法都不能完成这项工作。因为一切都唾手可得，人们必须小心选择自己的入手点，通常情况下，大家默认接受的前提就是罪魁祸首。我们在这狡猾的概念之域中探索，非常依赖思考工具的使用，那些思考工具设计出来就是为了帮人厘清头绪、揭示出不同思路的前景。

这些思考工具很少给我们提供一个稳固的支点，它不提供一条指导未来研究的“公理”，而是提供了一个适当的选项，为未来研究的可能性做出了一些约束，我们可以在以后的研究中对它进行修改，如果理由充分，还可以干脆将它放弃。难怪有很多科学家对哲学一点儿兴趣也没有：在哲学中一切都唾手可得，却找不出能像存在银行里的财物那样保险的东西，就连为了连接“支点”而构建起来的那些复杂的论证网络也都只是暂时性的，没有什么可用于证实或证伪的经验作为明确的基础。于是这些科学家对哲学置之不理，继续自己的工作，但这也令他们错过了一些最重要也最有意思的问题。“别问我这些问题！也别跟我说答案！想要解决意识、自由意志、道德、意义和创造力的问题，现在还太早！”

不过很少有人能如此节制，近些年来，科学家们已经在这些他们原来唯恐避之不及的领域中掀起了一波又一波的“淘金热”。在好奇心以及名声的驱使下，他们开始琢磨“大问题”，做着做着就会发现在这些问题上想要前进一步有多难。必须承认，我最享受的一件乐事就是瞧着那些在几年前还对哲学极尽挖苦之能事的知名科学家[7]，跌跌撞撞、绞尽脑汁地想替全人类把这些问题摆平，结果却只从自己的老本行中推出一些肤浅的论断。要是他们转向我们这些哲学家，承认他们需要从我们这里获得一点帮助的话，我就更开心了。

在随后的第一部分中，我会分别介绍12个多用途的一般性思考工具。在后续的部分中我则不再依工具的种类，而是按照它们所处理的主题分类来介绍。我首先会谈一谈最基本的哲学主题：意义或者内涵。然后是进化、意识和自由意志。我介绍的工具是真正的“软件”，这些工具之于你的想象力，就像显微镜和望远镜之于你的肉眼。

我顺便也会介绍一些似是而非的工具，这些工具非但不帮我们照明，反而放了些烟雾弹。我必须给这些危险的工具起一个名字。得益于我的航海经验，我在航海术语里找到了一个恰到好处的措辞。很多水手喜欢用航海术语忽悠“旱鸭子”，什么左舷、右舷、舵轴、舵销、桅支

索、横撑杆、索耳、导缆孔，等等。我以前待过的船上流行过一种玩笑，我们编造这些航海术语的定义：罗经座（binnacle）指的是在罗盘里长大的海员，桅杆柄脚（mast tang）是一种在高处享用的柑橘汁，开口滑车（snatch block）是一招女子防身术，吊杆托架（boom crutch）是一种爆破式整形装置。吊杆托架其实是当航速减慢、吊杆收起时用到的一种装在吊杆上的可移动木制托架，但我在此之后总是忍不住把它与某个倒霉蛋的胳肢窝发出“砰”的一声巨响的画面联系到一起。因此我给那些失败的思考工具起了“吊杆托架”这个名字。这些工具只是看起来增进了我们的理解，但实际上它们在散播黑暗、混乱而非光明。在后面的部分中散落着各种“吊杆托架”，它们被贴上了适当的警告标签，并配有一些令人遗憾的事例。

在本书结尾处，我会把我的反思继续向前推进，想一想做个哲学家是一种怎样的体验，万一你们对这个问题感兴趣呢？所以你们的丹尼特大叔，给每一位可能已对哲学探究世界的方式有所体悟，同时又怀疑自己是否适合做哲学研究的人，准备了一点儿忠告。

第一部分

通用思考工具

INTUITION PUMPS
AND OTHER
TOOLS FOR THINKING

这本书里介绍的大多数思考工具都具有很强的针对性，将它们制作出来就是为了处理特定的论题，甚至是为了处理特定论题中的特定争论的。不过在我们把目光转向这些特殊的直觉泵之前，先来看看那些通用的思考工具，这些构想、这些思维训练已经在种种情境中证明过自己的价值了。

有人说：“宁可什么都不信，也比信了一句谎言要强。”说这话的

人不过是在表现他有多么害怕成为一个傻瓜……这就如同一位将军告诫他的士兵们，即使永远不上战场也比冒受点小伤的风险强。但如果抱着这样的态度，我们就既不能战胜敌人也无法征服自然了。我们犯的错误真的没有那么严重、可怕。在这个无论人们多么小心谨慎也一定会犯错误的世界里，与极度神经质比起来，些许轻松的心态似乎对我们的健康更有利。

——威廉·詹姆斯，《信仰的意志》

一旦你下决心要测试一个定理，或者是说明某些观念，那么无论结果偏向哪一方，你都应该把结果发表出来。也许单发表某些结果，我们可以把论据粉饰得漂亮、堂皇，但事实上，我们一定要把正反结果都发表出来。

——理查德·费曼，《别闹了，费曼先生》

01

犯错儿
犯“好”错儿才有价值

科学家们常会问我，你们哲学家为什么要在哲学史的教学上花那么大的力气？化学家一般只掌握一些在我们读书时顺便积累下来的化学史基本知识，许多分子生物学家似乎也不关心1950年之前在生物学领域里发生过什么。我的回答是：哲学史实际上是记录了一大堆非常有智慧的人犯下一大堆非常有诱惑力的错误的历史，如果你不了解它，就注定会再次犯下那些倒霉的错误。

因此我们要向学生教授哲学史，如果科学家们轻率地忽视哲学，那只能后果自负喽。没有哪种科学不包含哲学，只是有些科学家并不反思其中潜在的哲学假设。有时候，最聪明又最幸运的科学家很灵巧地躲开了陷阱，可能他们属于“天生的哲学家”，也可能他们确实有自己以为的那么聪明。但他们只是稀有的例外。我并不是说专业的哲学家就不会再犯那些错误，甚至他们还会为那些旧的错误辩护。毕竟，如果问题很简单，也就不值得哲学家们费心思索了。

有时，你只是不想冒险犯错儿，如果能获得一些具体、清楚、确定的东西，犯点错儿也无妨。犯错儿是取得进步的关键。当然，有些时候真的一点错儿也不能犯，去问问外科医生和飞行员就知道了。但是没有多少人知道，有时候犯错儿才是我们唯一的出路。很多在竞争激烈的大学中读书的学生都以自己不犯一丁点儿错为荣，他们认为，或者被人诱导着认为，这就是自己能比其他同学走得更远的原因。我发现我常常需要鼓励他们培养犯错儿的习惯，因为犯错儿才是最好的学习机会。学生们有时会患上“写作障碍”，浪费时间绝望地在起跑线上徘徊。“脱口而出！”我敦促他们。这时他们才能在纸上写点儿什么。

哲学家是犯错儿的专家。我知道这听上去像个拙劣的玩笑，但是请听我说完。当其他学科的研究者专心地为定义好的问题寻找答案时，我们哲学家还在绞尽脑汁地找寻那些含混不清、极其错误的东西，甚至无

法确定正确的问题，更不要说正确答案了。提出错误的问题会令接下来的全部研究产生误入歧途的风险。只要发现了上述情况，你就可以认定自己做的是哲学家的工作没错！哲学是那种需要你首先想清楚应该问什么问题才能继续研究的学科，每个哲学领域中的研究都是如此。有些人讨厌这种事。他们宁愿自己需要回答的是现成的问题，这些问题做工精良，洗净了、熨平了，就等着他们来回答。这些人可以去做物理学、数学、历史学或者生物学研究。当然，有很多工作可以选择。我们哲学家喜欢研究的就是那种必须先把它们厘清，然后才能给出答案的问题。这个工作不是人人都能干的。不过，试试看吧，可能你会爱上它。

在这本书中，我时刻准备痛斥别人所犯的错误，可是我向你保证，我自己就是一位经验丰富的错误制造者。我犯过一些很棒的错儿，希望还能再犯更多。我写这本书的一个目标就是帮助你们犯“好”错儿，也就是那些能增进大家理解的错误。

人们总说：“先理论，后实践。”但犯错儿不只是学习的机会；从某种意义上讲，它还是人们学习或者做出真正创新的唯一机会。必须先做一个学习的人，然后才能学到东西。除非出现奇迹，否则人们只能通过两种方式真正学到东西：要么自己进步，要么接受进步者的安排和改造。生物的进化就是以大量残酷的试错为基础展开的，如果没有错误，即便尝试也将一无所得。就像美国著名作家戈尔·维达尔（Gore Vidal）说过的那样：“成功是不够的，必须还有人失败。”尝试可以是盲目的，也可以有着长远的考虑。对于已经拥有知识，但还没找到问题的答案的你来说，可以冒一次险，冒一次有着长远考虑的险。你可以三思而后行，从而在一开始就被已知所引导。你不必瞎猜，不过也不要瞧不起“瞎猜”哟；在通过“瞎猜”产生的所有造物中，最奇妙的那个就是——你！

就像我所写的其他书一样，本书的主题之一也是进化论，原因很简单：不仅对于生命而言，对于知识、学习和理解而言，进化问题也是根本且核心的。如果没有合理而完备的进化论知识，想要理解意义与观念的世界、理解自由意志与道德、理解艺术与科学甚至哲学本身，我们都会感到力不从心。

稍后我们将看到一些在处理更有挑战的进化论问题时所用的思考工具，不过现在我们只需要一些基本的认识。进化无法预测无目的突变通往何方，那是随机发生的DNA复制错误。大多数这种复制错误都没有结果，因为它们不会被读到，就像无关紧要的草稿纸，你没有也不准备把它们交给老师打分。一个物种的DNA就像一份打造新躯体的配方，大多

数DNA从未参与建构身体的过程，因此它们常常被称为“垃圾DNA”。对于那些被读取出来并在生物体的生长过程中发挥作用的DNA序列来说，大多数突变都是有害的，事实上很多突变是可以立即致死的。因为大多数突变基因的“性状表达”是有害的，自然选择实际上将突变率始终保持在一个很低的水平上。你们身体的每一个细胞中都有着非常非常好的“复制机”。例如，你的体内有一万亿个细胞，每个细胞中都包含你的基因组的一个完美或几乎完美的复本，长度超过了30亿个字符，这是自你父母的精卵细胞结合时就为你的存在配好的配方。幸运的是，复制并不能做到绝对完美，因为如果它做到了，它更新的源头将会枯萎，进化终将停止。进化历程中的这些微小的瑕疵与“缺陷”反而是生命世界中一切美妙、复杂的设计的源头。我忍不住要说：如果真的存在什么称得上“原罪”的东西，那一定就是这些复制失误了。

犯错儿的要诀在于不去隐藏它们，尤其不能自我欺骗。与其在我们犯错儿后扭脸否认，不如成为自己错误的鉴赏家，打心眼儿里把它们视作一种艺术品，其实在某种意义上它们就是艺术品。无论面对什么错误，我们的反应常常是：“好吧，我不会再犯这个错儿了。”而实际上自然选择并不会有这种想法，它只是在那些犯了错儿的家伙们能够繁衍后代之前把它们从地球上抹去；自然选择也不会犯两次同样的错，至少这比较罕见。能够学习的动物的大脑中有某种相同的选择功能，比如能够学会不发出噪音、躲避电线、选择自己的食物。B. F.斯金纳（B. F. Skinner）等行为主义者把这种需要称为“强化”学习，带来不良结果的反应不会被强化，更不会“灭绝”。我们人类把这个功能提升到了一个更迅速、更高效的水平上。实际上我们可以反思自己的想法，也可以反思自己刚刚做过的事：“好吧，我不会再犯这个错儿了！”当反思时，我们会直接面对每一个犯错者都必须解决的问题：到底错在哪儿了？究竟我刚才做了什么，使自己陷入了这个窘境？关键在于利用错误当中那些特殊的细节，这样你才能在下次尝试时对错误有所察觉，以免重蹈覆辙。

我们应该都听过那种可怜的托词：“好吧，不过它当时看起来是个好主意！”这种托词似乎是一种愚蠢的标志，一种傻瓜似的反省，但实际上我们应该像欣赏智慧的基石一样欣赏它。任何存在，任何行动者，只要能够真心地说：“好吧，不过它当时看起来是个好主意！”那他就已经站在了通向卓越的门口。我们人类为自己的智慧而自豪，其标志之一就是我们能够记住自己之前的想法，并进行反思，反思我们怎么会有这样的想法，之前为什么会被它所吸引，以及哪里出了问题。就我所知，没有任何证据表明地球上的其他物种能够想到这一点。如果它们能的话，那

就几乎像我们一样聪明了。

所以，当你犯了错儿时，应该学会深吸一口气，咬紧牙关，然后尽你所能地冷静、毫不留情地检查你记住的错误。做到这一点其实并不容易。人类对错误的自然反应是尴尬与愤怒（我们永远是对自己最生气），因此你必须努力克服这些情绪反应。尽管这是个有些古怪的实践，但是试着学会品尝你的错误吧，为将你引入歧途的怪癖被揭露出来而感到开心吧。当你把曾经所犯的错误中的所有养分吸干之后，就可以开心地把它们抛在脑后，向下一个伟大的机会前进了。不过这样做还不够：你应该积极地寻找机会去犯伟大的错误，只有这样，你才能从中吸取教训。

简单地说，这是一种我们在小学时就已经掌握的技能。回想一下你第一次做长除法时的感受：你看着两个无法推算的大数，不得不琢磨从哪儿入手，那是多么陌生、多么令你望而生畏的感觉啊。被除数是除数的六倍、七倍还是八倍呢？谁知道呢？你不必知道，只需要试一试，随便挑一个你喜欢的数，算算结果如何。我记得当老师告诉我应该从“猜数”开始的时候，我几乎震惊了。这还是“数学”吗？我怎么能在如此严肃的活动中玩猜数游戏啊？但是最终我还是像大家一样学会了欣赏这种策略之美。如果发现选取的数小了，你就换个大一点的重新开始，如果发现选大了，那就换个小点的数呗。长除法的好处在于，你总是可以算个结果出来，即使你第一次选的时候特别笨，那也只是多花点时间而已。

或多或少熟练地做出猜测，算出可能的结果，利用结果为下一步工作做出修正，这样的一般方法已经得到了广泛的应用。这个方法的关键在于，所犯的错误要足够清楚、精确，使得出的结果足够确定。在全球定位系统（GPS）发明之前，水手们在海上为船只定位的第一步工作就是做出猜测，他们会猜一个确切的经纬度，然后假设所猜的恰好就是他们实际所处的位置，按照这个位置推算出太阳此时在天空中“应处”的确切位置。当使用这一方法时，他们并不期待一击即中。没有这个必要。因为他们接下来还会测量此时太阳的实际仰角，并且比较前后两个数据。通过一番简单的计算之后，他们就会知道应该对最初的猜测朝哪个方向做出多少修正了。[8]使用这种方法时，一开始猜得很好当然是有帮助的，但必定会出差错也没有什么关系，重要的是错误要犯得清楚、明确，这样才有东西供我们去校正。GPS导航仪也是使用同样的“猜测–修正法”来确定它与人造卫星之间的相对位置的。

当然，你面对的问题越复杂，分析起来就越困难。这在人工智能

（AI）研究中被称为赞誉分配问题或责备分配问题。搞清楚什么应该获得赞誉，什么应该被责备，是人工智能研究中最棘手的问题之一，而这也是自然选择所面对的问题。

在讲完其一生跌宕起伏或平凡无奇的故事之后，地球上的每个生物终将死去。大自然如何能够看穿所有细节的迷雾，搞清楚哪些是应当鼓励的、让它们继续繁衍生息下去的正面因素，哪些是应当惩罚的、让它们无后而终的负面因素，进而做出自然选择呢？难道我们祖先的亲戚们真的是因为眼皮长错了形状才灭绝的吗？如果并非如此的话，如何用自然选择的过程来解释我们的眼皮为什么恰好长成这样？

这个问题的一部分答案其实我们已经能够猜到了，像那句老话说的："如果东西没坏，何必要修。"把原来一切保守的设计方案都留下吧，冒险的时候先准备一张安全网。自然选择会自动地把那些"到现在为止"管用的东西都保留下来，同时无所畏惧地开展或大或小的创新。只是那些巨大创新几乎总会立即导致死亡。这是一种巨大的浪费，不过没人能去计算。我们的眼皮早在人类存在以前，或者早在灵长类甚至哺乳类动物存在以前就被自然选择所造就了。它们用了超过一亿年的时间演变成了现在的形状，只在最后六百万年的时间里做了些小修小补，那时正值我们和黑猩猩、倭黑猩猩的共同祖先生活的时代。这个问题的另一部分答案是：自然选择包含了大量的个案，即使是统计上的微小优势也会自动积累。答案的其余部分则过于技术化，超越了我们现在所做的初级讨论的范围。

卡牌魔术师有一个效果惊人的绝招，起码最优秀的卡牌魔术师会这一招。我不想因揭秘一个魔术而激起魔术师们的公愤，因此我要讲的只是技巧背后的一般原理。一个好的卡牌魔术师明白，其实很多招数并不总是管用，反而要依赖运气。有些魔术成功的概率可能只有千分之一！这应该已经不能被称为一种魔术了。

我要讲的那一招是这样的：在告诉观众们你要表演魔术时，千万不要告诉他们是哪种魔术，你首先做那个成功率只有千分之一的魔术。当然，这个魔术几乎肯定要失败，这时你应该马上不动声色地做下一个魔术，比如一个成功率有百分之一的魔术，如果这也不成功（通常也会如此），那你就优雅地、悄无声息地开始第三个魔术，它的成功率可能有十分之一。现在最好也把第四个魔术准备好，假设它的成功率是百分之五十。如果这四个魔术都失败了（这种情况一般不太可能发生），你就表演一个不会给观众留下太多印象，但一定会成功的魔术吧。如果在表

演时你总要依赖最后的安全网才能脱身，那可真是太不走运了。不过，只要高难度的魔术做成功了一次，观众们就会惊呆的。“这不可能！你怎么可能知道这是我的牌？”啊哈！你并不知道，是你聪明的瞎猜法给了你回报。通过把所有“错误”不露马脚地隐藏起来，你就创造了一个“奇迹”。

进化也是按照这种方式进行的：所有愚蠢的错误都慢慢消失了，因此我们看到的只是一系列惊人的成功。你们知道吗，有史以来绝大多数的生物都没有后代，但我们的祖先中却没人遭此厄运。想想从你的祖先一直延续到你的这一连串多姿多彩的人生吧。

舞台魔术与科学有一个很大的区别在于，魔术师应当尽一切办法向观众隐瞒他们的错误，而科学家所犯的错误应该公开。你把自己的错误公布出来，才能让别人从中吸取教训。这样一来，你也能从别人的经验当中获益，避免自己“一条道走到黑”。物理学家沃尔夫冈·泡利（Wolfgang Pauli）在批评他同事的工作时说过一句名言，他鄙视地说：“你的工作连错误都算不上。”一个给人以批评空间的清楚、明确的错误比一团糨糊要好多了。顺便说一句，这也是我们人类比其他物种聪明得多的一个原因。与其说我们胜在脑容量更大、大脑功能更强，或者具有反思自己过往错误的能力，不如说我们胜在可以相互分享个人在试错的历史中所获得的经验。(9)

我很惊讶有那么多真正的聪明人并不明白，你可以把你所犯的严重错误公布出来并且表现得一点儿也不在乎。我知道，杰出的学者在否认他们犯过的错误时可以变得多么强词夺理。显然他们从来没有认识到，就算承认“哎呀，你是对的，我想我犯了一个错误”，也没人会吃了你。实际上，人们“乐意”看到别人承认自己的错误，也都乐意指出别人的错误。看到你承认了错误，心胸宽广的人会感谢你给了他们帮忙的机会，并给予你回报；而心胸狭窄的人则乐呵呵地看你出洋相。无所谓喽！反正所有人都是赢家。

当然，一般人们并不乐意纠正别人犯的“愚蠢”错误。你必须犯一些值得纠正的错误，无论如何要犯得有原创性，例如有的错误需要人在犯这个错的同时搭建一座有挑战性的思维金字塔，这就犯得很有特色，就像我们在卡牌魔术中看到的那样。谨慎地利用别人的批评，能把你从孤立无援的境地中解救出来。还有一个意想不到的好处是：如果你是一个勇于冒险的人，人们就乐意纠正你偶一为之的愚蠢错误，因为这些错误能够显得你不是那么特别，显得你也像我们其他人一样笨手笨脚的。

我认识一些特别谨小慎微的哲学家，他们在工作中从来没有犯过明显的错误。这些人往往不愿意干得太多，他们做的那一点工作也是非常原始的，对于这些人来说，除此之外的工作都太冒险了。指出别人的错误是他们的专长，这可以说是一项极有价值的服务了，但是只要他们犯了一点小错，别人却从来不会一笑而过。可以说，这是很不幸的，他们最好的工作也常常会因此蒙上一层阴影，比如被人们忽视，或被更大胆的思想家们驱动的当下的潮流所淹没。在第76章中，我们将会看到，虽然总的来说敢于犯错儿是件好事，但它也有不幸的一面。给你一个“元”忠告：对任何忠告都不要太认真哦！

02

归谬法
发现错误命题的妙招

它是理性研究的撬棍，它是贯彻一致性的伟大杠杆，它就是归谬法。顾名思义：归（推理）于荒谬。你把一个推论中的一些命题或假设拿过来，看看有没有前后矛盾，或者是否隐含着荒谬之处。要是有，必须丢掉出问题的命题，或者重新推论。我们一直在这么理所当然地做着："如果那是一头熊，那么熊就是头上长犄角的！""他赶不上晚饭了，除非他像超人一样能飞。"

当一个命题有着复杂的理论争议时，撬棍通常舞得飞起，但在这里我们却很难区分公允的批评与嘲讽式的反驳。你的对手真的会笨到相信我们用几步轻巧的归谬就能推翻的命题吗？我以前给一个学生的论文打分时，发现了一处拼写错误，他把"parity"拼成了"parody"，创造了"以蠢类推"这个美妙词组[(10)]，我认为这是一个比"归谬法"还顺嘴的叫法，它在粗糙而混乱的科学、哲学的争论中得到了普遍使用。

记得几年前，我在麻省理工学院参加了由语言学家诺姆·乔姆斯基（Noam Chomsky）和哲学家杰里·福多尔（Jerry Fodor）主持的研讨会。在会上，从外地赶来参会的认知科学家如果没能讲得令观众们满意，就会收到滑稽的辩驳。有一天，人工智能理论家、认知心理学家罗杰·尚克（Roger Schank）成了大家的"眼中钉"，他是耶鲁大学人工智能实验室的主管，在乔姆斯基的口中，他简直蠢得冒烟。我很了解尚克和他的工作，虽然我自己并不完全赞同，可是我也难以苟同乔姆斯基的看法，于是我举起手来跟乔姆斯基说，他可能漏掉了尚克的立场中的一些微妙之处。"哦，不对，他就是那么想的！"乔姆斯基咯咯地笑着否定了我，然后就继续拿尚克的工作开涮，继续做这件最令听众们感到开心的事去了。几分钟以后，我实在忍不住又打岔道："我不得不说，你批评的那种观点简直是太荒谬了。"乔姆斯基笑着点点头，我接着说道："那我就想问问了，咱们干吗还要浪费时间来批评这种垃圾呀？"这桶冷水泼

得！

我自己“归谬”别人的观点时又干得怎么样呢？有没有公允一点呢？下面举几个例子，你们来评判一下。在威尼斯的一次会议上，我曾和法国神经科学家让-皮埃尔·尚热（Jean-Pierre Changeux）一起与澳大利亚神经科学家约翰·埃克尔斯（John Eccles）和奥地利籍英国哲学家卡尔·波普（Karl Popper）就意识和大脑的问题进行辩论。尚热与我是唯物论者，坚持心灵即大脑，而波普与埃克尔斯是二元论者，声称心灵不像大脑那样是物质的，而是属于某种能与大脑发生作用的第二实体。埃克尔斯在很多年前因发现突触而荣获诺贝尔奖，突触是神经元之间微小的间隙，每天有上万亿的谷氨酸分子和其他神经递质、神经调质通过这些间隙。根据埃克尔斯的说法，大脑是以上万亿的突触为琴键做成的强大的管风琴。这位虔诚的天主教徒还认为，非物质的心灵，也就是不朽的灵魂以某种方式在量子层面上推动着谷氨酸分子，从而弹奏着突触琴键，真是激动人心啊！他说：“忘了所有那些关于神经网络之类的理论的讨论吧，都是些不相干的废话。心灵就在谷氨酸中！”轮到我发言时，我说我想确认一下自己是否理解了他的观点。如果心灵就在谷氨酸中的话，我把一碗谷氨酸倒进下水道，是不是就是杀人了呢？“唔，”他有点吃惊地说，“这个问题不好回答呀，不是吗？”[11]

你可能会想，除了都获得过诺贝尔奖之外，这位虔诚的天主教徒约翰·埃克尔斯与无神论的唯物论者弗朗西斯·克里克（Francis Crick）之间不会有什么共同点了吧。其实，他们在关于意识问题的看法上都有些过于简单化了。科学圈之外的人大多不了解“过度简化”在科学中是一种多么神奇的东西；科学家们可以用这种“几乎”正确的工作方式绕开令人讨厌的复杂性，把纷乱的细节搁置一旁。

可以说，在科学史当中把“过度简化”用得最成功的要数克里克和詹姆斯·沃森（James Watson）在发现DNA结构时玩的那套“迂回战术”了，在他们公布这个发现时，莱纳斯·鲍林（Linus Pauling）等人还在为厘清所有的细节而埋头苦干呢。克里克之所以赞同做这种大胆的尝试，在于他想着万一能够把问题一举攻克呢，但是这种做法显然并不总会奏效。我曾经在拉霍亚参加过克里克举办的一个著名的茶话会，这让我有机会证明这一点。这种午后讨论会是他们实验室组织的非正式会议，客人们也可以参与一般性的讨论并提出问题。恰好在这次讨论会上，克里克做了一个大胆的声明：“我们最近已经发现，大脑皮层V4区的神经元与颜色‘相关’，它能对不同的颜色产生不同的反应。”然后他就

提出了一种非常简单化的设想：以红色为例，人对红色的经验意识“就是”视网膜区域的红敏神经元的活动。嗯……我听着有些纳闷儿：“你是说，如果我们把一些红光敏感神经元放到培养皿中并让它们保持活性，再用微电极刺激它们，那么在培养皿中也会出现对红色的意识吗？”

有一个可以对付归谬法的招数，那就是迎难而上、死不松口。自从澳大利亚哲学家J．J．C.斯马特（J．J．C．Smart）为了坚持他的伦理理论，惊世骇俗地说出“对呀，把一个无辜的人绞死有时候就是正确的”这种话以后，我就给这招起了个名字，叫“我比你斯马特”。[(12)]克里克决定他要比我更“斯马特”了，他说：“对呀，会有一种独立的对红色的意识！”谁对红色的意识呢？他没有说。虽然不久之后，他对这个设想做出了改进，但是在他与神经科学家克里斯托夫·科赫（Christof Koch）一起寻找他们所谓的NCC（意识的神经相关物）时，克里克却表示他从未真正放弃过对这一想法的忠诚。

不过，或许我在另一个场合里对“‘培养皿中能有意识出现’这种想法到底荒唐在哪儿”这个问题的归谬更好。物理学家兼数学家罗杰·彭罗斯（Roger Penrose）和麻醉学家斯图尔特·哈梅罗夫（Stuart Hameroff）一起建构了一种意识理论，他们认为意识并不是谷氨酸，而是神经元微管中的量子效应。微管是一种管状的蛋白链，它们作为支架与导管存在于所有细胞的细胞质中，而不只在神经元中。

在第二届图森会议，即第二届国际意识科学大会上，当哈梅罗夫表述了这种观点之后，我在观众席上提了这么一个问题：“斯图尔特，你是一个麻醉学家吧？你有没有做过一种戏剧性的手术，就是那种给断手或者断胳膊的伤者做的再植手术？”他说他没做过，但知道这种手术。“那么好了，斯图尔特，如果我哪里搞错了就请你提醒我。假如你是一个再植手术的麻醉师的话，按照你的理论，你在道德上是不是有义务给那只断掉的手打麻药啊，它还躺在冰袋里呢，对不对？毕竟像神经系统中的其他微管一样，断手中的神经元微管还干着人家该干的事呢，所以那只手难道不会感受到剧痛吗？”此时斯图尔特脸上的表情告诉我，他从来没有想到过这种情况。无论是对红色的意识、疼痛的意识，或者随便什么意识，说它是归某种神经网络所有，或者把它归于无数神经元的协调活动，这种想法可能最初并不是很有吸引力，但是这些“归谬”的尝试却有可能帮助我们认识到，这种想法为什么应该被我们认真对待。

03

拉波波特法则

批评他人的正确方式

我们到底应该对对手的观点报以多大程度的宽容呢？如果对手所说的话中有非常明显的矛盾，那么你当然应当尽力把它指出来。如果他们的说法中潜藏着一些矛盾，你就应该小心翼翼地把它们揭露出来，然后再施以攻击。但是，我们在寻找对方说法中隐藏的矛盾时常常容易走过头，变成鸡蛋里挑骨头，变成彻头彻尾的嘲弄，变成“海事律师的行当”[13]，就像我们在上一节中看到的一些例子一样。这种认定对手的想法中一定藏有矛盾的冲动，给了你一个容易击中的靶子，怂恿你对别人的看法做出无情的解读。可是，这个容易击中的靶子通常与我们真正关心的问题八竿子打不着，你只是在浪费大家的时间和耐心而已，最多是给支持你的观点的人提供了一点消遣。

很多年前，社会心理学家兼博弈论专家阿纳托尔·拉波波特（Anatol Rapoport）曾提出过一套规则，我认为对于那种讽刺对手的恶习，这套规则是最好的解毒剂。在罗伯特·阿克塞尔罗德（Robert Axelrod）的囚徒困境锦标赛中胜出的策略“一报还一报”就是拉波波特的发明。[14]

拉波波特法则：如何撰写一篇成功的批评性评论。

① 你应该非常清楚、生动、不偏不倚地重述对手的想法，使得你的对手说：“谢谢你，我刚才要是像你这么表述就好了。”

② 你应该把对方观点中你所同意的部分都列出来，尤其是那种并非被人们广泛接受了的观点。

③ 你应该提到那些从你对手那里学到的东西。

④ 只有完成了以上三点，你才能说一句反驳或批评的话。[15]

遵循此法则的一个直接的好处就是，你的对手会因此乐于倾听你的

批评。因为他们看到你已经站在他们的立场上理解了他们的想法，同时你也展示出了良好的判断力，在一些重要的事情上和他们达成了共识，甚至在有些地方还被他们说服了。

遵循拉波波特法则时总是令人倍感挣扎，起码对我而言是这样。说白了，有些对手不值得我们如此尊重。我不得不承认，把他们串起来烤了真是一件乐事。但当我们用了这个方法而且使用得当的时候，总是后果喜人。在我的《自由的进化》（*Freedom Evolves*, 2003）这本书中，我特别努力地试图公正对待罗伯特·凯恩（Robert Kane）所持的那种“不相容论”，这是一种我非常不赞同的自由意志观。在我把那一章的初稿发给凯恩之后，他给我回了信，而这封信也被我珍藏了起来：

> ……你的文章真的让我非常欣赏，尽管我们的看法不同。你对我的观点的评论非常全面而公正，远远超出通常的批评。你考虑到了我的观点的复杂性以及我努力解决那个棘手问题时所抱有的严肃态度，而没有将它们一棍子打倒。为此，也为了你细致的批评，我非常感谢你。

我也遵循着拉波波特法则批评过另外一些人，他们的反应就没有这样亲切了。有时候，批评看起来越公正，接受起来反而越难。你应该时刻提醒自己：对于一位作者来说，如果你勇敢地为他的观点寻找根据，但什么也没找到的话，可能会比气呼呼地批评他的观点来得更有杀伤力。所以，我向大家推荐拉波波特法则。

04

史特金定律
不要把时间浪费在无意义的事情上

在1953年9月，科幻作家泰德·史特金（Ted Sturgeon）在费城举办的世界科幻大会中的发言中提到：

> 当人们谈起悬疑小说时，会提到《马耳他之鹰》（*The Maltese Falcon*）和《长眠不醒》（*The Big Sleep*）。当人们谈起西部小说时，会想到《西部新天地》（*The Way West*）与《原野奇侠》（*Shane*）。可当人们谈起科幻小说时却只会把它叫作“什么巴克·罗杰斯（Buck Rogers）之类的”，还说，“百分之九十的科幻小说都很‘渣’（crud）”。好吧，人家说的没错。百分之九十的科幻小说确实很渣。但是，任何事物当中的百分之九十都是渣滓，而其中那不是渣滓的百分之十才是有意义的。我敢说，那百分之十的科幻小说可以与人类从古至今写就的任何作品相媲美。

人们一般会把史特金定律表达得更粗俗一些：“任何事物当中的百分之九十都是垃圾（crap）。”百分之九十的分子生物学实验、百分之九十的诗歌、百分之九十的哲学书、数学系里百分之九十的同行评议文章，等等，都是垃圾。真的吗？嗯，这么说可能有点极端了，不过我们必须承认，在每个领域中都有大量平庸的工作。有些个“爆脾气”觉得垃圾差不多得占百分之九十九……我们就不要跟着他们一块儿义愤填膺了。

从这一观察中我们应该学到的是：当你准备批评某一领域、某一流派、某一学科或者某一艺术形式的时候，请别在嘲笑垃圾上浪费自己和我们大家的时间！要么就追着优秀的东西批，要么就别干这活儿。可理论家们却常常罔顾这一忠告，被他们诋毁过的可太多了：分析哲学、进化心理学、社会学、文化人类学、宏观经济学、整形手术、即兴表演、情景喜剧、哲理神学、推拿疗法，无奇不有。不论你要评论什么领域，

让我们从一开始就规定好：每个领域都有一大堆愚蠢、平庸、糟糕透顶的东西存在。然后，为了不浪费你的时间和大家的耐心，请确保你关注的是你能找到的该领域最好的东西，比如获奖之作、杰出成果、大家都赞不绝口的精品，而非糟粕。

请记住，你不是个喜剧演员，你的主要目的不是讲个笑话逗人一乐，不是跟我们分享夸张的一幕。这一点与拉波波特法则紧密相关。我发现，当我们批评的是哲学家时尤其应当牢记这一点。从最伟大、最敏锐的古希腊先贤，到昨日的智性英雄——比如这四位非常不同的思想家：伯特兰·罗素、路德维希·维特根斯坦、约翰·杜威、让–保罗·萨特——经过些许巧妙扭曲之后，他们最好的学说、最精妙的分析也会显得又臭又长，变成彻头彻尾的蠢话。别这样好不好？这样只会丢你自己的脸。

05

奥卡姆剃刀
如无必要，勿增实体

这个思考工具得名于奥卡姆的威廉（William of Ockham），一位14世纪的逻辑学家、哲学家，不过它实际上是一条古老的经验法则。它的拉丁语是“lex parsimoniae”，即节约律。在英文中人们常常用格言“如无必要，勿增实体”（Do not multiply entities beyond necessity）来表达。这一想法非常简单：如果你能找到一个更简单的理论，包含更少的因素、更少的实体，而它也一样可以很好地解释现象，就不要再构造一个大而无当的复杂理论了。比如，如果“暴露在极端寒冷的空气中”可以解释关于冻伤的一切症状，那就别再假定还存在未被发现的“雪细菌”或者“北极微生物”了。开普勒定律为行星的轨道给出了解释，所以我们不必假设地表下还藏着一个在控制台前驾驶行星的驾驶员。话讲到这儿，应该没人会有疑义，可当我们把这一原则扩展开去，就不一定总会获得认可了。

康维·劳埃德·摩根（Conwy Lloyd Morgan）是19世纪的英国心理学家，他把“奥卡姆剃刀”应用到了“动物是否拥有心智”这一问题中。摩根的节约律告诉我们，如果动物的行为可以用更简单的术语说明的话，就不要把高贵的心灵归于昆虫、鱼类，甚至海豚、狗和猫等哺乳动物。他说：

> 如果动物的行为能用刻画心理进化、心理发展的较低层次的术语来解释，就决不用刻画高级心理过程的术语来解释。（Morgan, 1894, p. 128）

扩展开来，这就相当于让我们把动物甚至人类看作是只有大脑却没有心灵的存在。正如我们以后会谈到的，人们考虑心灵这一主题时遇到的压力，不会因一种绝对禁令而消除。

很多人都想用“奥卡姆剃刀”把本来盘根错节的问题剃个干净，其中最令人印象深刻的努力是，有人声称上帝是宇宙的创造者，因为这种假定更简单、更节约。假定超自然的不可理解的东西存在怎么就成了更节约的做法了？我觉得这是一种极大的浪费，我们或许能找到一种聪明的办法来反驳它。我不想做过多的论证；我只想说，说到底奥卡姆剃刀只是一种经验法则、一条有益的建议。把它当作一种形而上学的原则或者理性的基本条件，期望用它来一举证明上帝存在或不存在，就简直太荒谬了。这就像当你反驳一个量子力学的定理时说，那条定理违反了格言“不要把所有鸡蛋放到同一个篮子里”。

还有一些思想家把奥卡姆剃刀用得更加登峰造极，比如用它来否定时间、物质、数字、洞、美元、软件等种种事物的存在。古希腊哲学家巴门尼德（Parmenides）是最早的“吝啬鬼”思想家，写在他的存在目录中的东西实在是少得不能再少了。就像我的一个学生在一次测验中写的：“巴门尼德就是那个说‘只有一物存在，而我也不是它’的人。”虽然我讨厌这么说，但这似乎就是巴门尼德想要告诉我们的。当然，他的思想在翻译的过程中会丢失一些东西。我们哲学家习惯于认真对待巴门尼德这样的思想，就是因为我们永远也不知道什么时候能说一个想法是“疯狂”的，不知道这会不会是一次不正当、不明智的判断，那个思想会不会成了我们失败的想象力的受害者。

06

奥卡姆扫把

有意隐瞒对自己不利的证据

分子生物学家西德尼·布伦纳（Sydney Brenner）最近发明了一种“奥卡姆剃刀”的新玩法，非常有趣，他称其为“奥卡姆扫把”，指某些理论的辩护者们本着智性上最大的不诚实把对自己不利的事实往地毯下面扫。这是我们遇到的第一个“吊杆托架”，第一个“反思想”工具，你们可要擦亮眼睛喔。宣传家们用这招来做群众工作时尤其阴险，因为就像福尔摩斯侦探小说中著名的“半夜狗不叫”的故事一样，只有专家才能注意到“缺席”的事实，而事实已经被“奥卡姆扫把”从现场扫走了。比如，创世论者就常常省略掉大量令他们的“理论”感到尴尬的棘手事实。对于一个不是生物学家的人来说，创世论者精心打造的论证可能是非常有说服力的，因为他们看不出什么“被扫走了”。

我们到底该怎样提防那些看不出的东西呢？向专家们求助。美国哲学家斯蒂芬·C.迈耶（Stephen C. Meyer）在他所著的《细胞里的签名》（*Signature in the Cell*, 2009）一书中试图系统性地证明生命不可能有一个自然的（即非超自然的）起源，并对所有持自然起源观点的理论模型做了一番纵览，以显示它们都是多么地无药可救，即使是一个相当博学的读者也可能会认为这番纵览看起来是公正、详尽的。

迈耶的这本书写得真是打动人心，就连著名哲学家托马斯·内格尔（Thomas Nagel）都在2009年11月给世界上最具影响力的书评刊物之一《泰晤士报文学副刊》写了一篇有关它的书评，说那本书是他心目中的年度最佳！在他那篇充满溢美之词的文章发表以后，我曾与他通信，内格尔在信中信心满满地表示自己对生命起源研究的历史有着相当多的了解，足以让他信任自己的判断。就像他于2010年1月1日写给《泰晤士报文学副刊》的一封信中所说：“在我看来，迈耶在真诚地写作这本书。”要是内格尔咨询了研究生命起源的科学家们的意见的话，他就能够发现迈耶是怎样利用“奥卡姆扫把”把那些对自己不利的事实扫出视野之

外的了，或许他还会失望地发现，迈耶并没有像对待他那样给那些专家们寄去样书，或在出版前询问他们的意见。发现他所赞美的这本书竟然是“秘密操作”，可能会动摇他对自己判断的信心吧，也可能不会。我们都知道，科研机构有时会不公正地压制那些叛逆的声音，所以有可能，仅仅是可能，迈耶别无选择，只能“发动偷袭”。但是内格尔在站出来替迈耶捧场之前还是应该谨慎地研究一下。确实，在关于生命起源的研究中科学家们并没有达成一致，也并不存在一个完全正确的理论，但这并不意味着我们没有合适的候选，这恰恰是因为竞争者太多而不是太少。

阴谋论者都是使用“奥卡姆扫把”的大师，你可以在网上做一个有趣的练习：找到某个新鲜出炉的阴谋论，看看你作为该领域的外行，能不能在搜索专家对它进行反驳之前找到它的错误。布伦纳发明这个术语的时候并不是在谈神创论或者阴谋论，他想说的是，即使是一位严肃的科学工作者有时也难以抵抗对某些数据的“忽视”，尤其是当那些数据会伤害到他们心爱的理论时。但无论如何，这都是一个需要抵御的诱惑。

07

外行做媒

既不“过少解释”，又不“惹恼行家”

我已经推荐这项技术很多年了，它可以很好地防止人们在不经意间挥动“奥卡姆扫把”，虽然它已经通过了多次考验，但次数还是没能达到我的期望。和我讲过的其他思考工具不同，这项技术的应用需要花费一定的时间和金钱。所以我希望有人能够把它大力推行一番，并报告推行的结果。在这儿我决定谈谈它，因为这里要谈的是人们使用其他一般性思考工具时也会面临的“沟通问题”。

不只是哲学，在很多领域中都有一些似乎永无止境的争论，造成这一状况的部分原因是人为的：因为争论的双方各说各话，没有为有效的交流做出必要的努力。人的火气一上来，蔑视和嘲讽也就接踵而至。这时就连看热闹的人也会纷纷站队，即使他们还没有完全搞懂人家谈的是什么。

场面可能会变得很难看，而造成它的原因可能非常直接。当专家与专家对话时，不论他们的研究领域是否相同，总会错误地“过少解释”。理由很简单：跟一位专家过度解释某件事是一种严重的侮辱，就好比对他说：“用不用我教你怎么拼写啊？”没人愿意侮辱一位专家。所以，为了安全起见，人们宁愿错误地“过少解释”。其实在大多数情况下，人们并非有意为之，而且要完全不这么做几乎是不可能的。这实际上是个优点，因为自觉保持礼貌终归是一种好品格。但是很不幸，“假装听众对你所说的内容已经非常了解”这种高尚的品格有一个副作用：专家们常常自说自话。

这个毛病没有办法直接治：假如我们在研讨班或者学术大会上恳请与会的专家们不要“过少解释”他们的立场，这或许会得到他们热切的同意，但即使如此也是没用的。这只会让事情变得更糟，因为这时人们反倒会对那些可能会侮辱到他人的话题更加敏感。

不过，有一种间接的方法相当有效：让专家们对一群充满好奇心的业余听众表达他们的观点（比如我们塔夫茨大学的这群聪明的本科生），同时让另外一些专家旁听。旁听的人不必偷听，我没有建议大家走“歪门邪道”。相反，参与这个活动的每个人都应该知道：这一实验的目的就是让参与者能够安心地用“大家”都能理解的术语说话。由于讲话者是在对着作为诱饵的听众讲话，他们“根本”不必担心自己会侮辱其他专家，因为他们的话不是对着那些专家说的。我猜他们也有可能会担心侮辱到本科生，但这是另外一个话题了。如果一切进展顺利，专家A给本科生们解释一个有争议的问题时，专家B也在听。听着听着，B可能就会眼前一亮。“原来‘这’就是你想说的呀！我终于明白了。”也可能B听的时候没有什么反应，轮到他给同样的一群本科生讲解时，专家A会感到眼前一亮。这个方法可能并不完美，但一般都会进展顺利，让每个人都受益。专家们消除了一些人为的误会，本科生们也接受了一流的教育。

这项实验我在塔夫茨大学开展过很多次，多谢校方的慷慨支持。我亲自挑选了不到12位本科生，我在向他们简述他们的角色时说：“不要接受任何你们听不懂的话。如果遇到任何在你们看来混乱或者模糊的地方，我期望你们举起手来，打断专家的话，提醒他们做出解释。”他们确实需要提前仔细阅读必要的读物，保证他们不完全是相关领域的菜鸟，而是充满好奇心的初学者。他们喜爱这种角色，当然了，这可是让大人物给他们开小灶啊。同时专家们也会发现，在这些条件的限制下，他们通常会找到一些前所未有的能更好地能解释自己观点的方式。这些专家常年被一群群同事、博士生、高年级研究生们“保护”着，有时，他们真的需要这种挑战。

08

跳出系统
打破惯性思维的好方法

我们很难发现别人使用了“奥卡姆扫把”，因为我们很难发现别人把哪些令他难堪的事实扫出了视野之外，但更难的是发现侯世达（Hofstadter，1979，1985）所说的“跳出系统”（jootsing）。跳出系统是一个重要的策略，不仅在科学和哲学领域，在艺术领域也是。创造力是一种让人梦寐以求却只有少数人才能够拥有的才具，它往往会以一种前所未有的、让人意想不到的方式打破孕育它的系统所设定的规则，也许是古典音乐的和谐体系，也许是十四行诗甚至打油诗中的格律与韵律，也许是某种艺术的审美“教规”、良好形式。又或者，它打破的是某些理论或研究领域中的假设与原则。有创造性不只意味着追求新奇的东西——因为新奇的东西谁都能得到，只要把材料随机排列一番就行了——难的是跳出一个既定的体系，并且跳得有道理。可当一个艺术传统真的达到了“怎样都行”的阶段时，那些想变得有创造力的人就会遇到一个问题：这里已经没有可以反抗的固有规则、没有可以打破的预期了，让创造力显得惊人并且有意义的背景都失去了，那我们还能颠覆什么呢？如果你想颠覆传统，最好先去了解它。所以外行或新手很少能够想出真正有创造性的东西。

你可以坐到钢琴前，试试能不能弹出一段既动听又新颖的旋律，很快你就会发现这有多难。你可以按下所有的琴键，你可以选择所有的组合；但除非你能找到一些可依靠的东西，找到一些风格、流派、模式来遵循，并且在此基础上做些改变，否则你弹出的只可能是噪音。同时，并非所有违反规则的行为都能有好的效果。我认识至少两个成功的演奏爵士乐的竖琴师，但如果你想靠敲击邦加鼓演奏贝多芬的名曲来博取名声的话，这可能就不是一个好计划了。这个道理对科学也是适用的：在任何理论的争论中都有很多未经证实的前提，但逐一找到它们的漏洞再一一否定掉这些前提，或许不是一种在科学和哲学中取得成功的好办

法。这就好像你要改编一段作曲家乔治·格什温（George Gershwin）的曲子，并希望自己改编得有价值，但却一次改动一个音符一样。祝你好运吧！突变几乎总是有害的。事情会比我描述的还要困难，但你有时候也会走运。

建议某人以跳出系统的方式前进，有点像是建议一个投资者低买高卖。是啊，当然，就应该这么干，但问题是到底该怎么干？请注意，对投资的建议并不“完全”是空洞、无益的，号召人们跳出系统会更有帮助，因为它让你对自己的目标看得更清晰了，正如每个人都知道“更多的钱”是什么意思。当你面对一个科学或哲学问题时，需要跳出的系统常常根深蒂固，以至于它像你所呼吸的空气一样令人不易察觉。一般来说，如果一个长期存在的争论似乎无法取得进展，双方都顽固地坚持自己的正确性，往往是因为他们都同意的某个东西其实是错误的。两方都认为这是明摆着的事，是不言而喻的。发现这些不可见的问题不是一个简单的任务，因为对论战的专家们来说是显而易见的东西，对于每个人来说也都是如此。所以你最好睁大眼睛去找出被双方共同默认的错误假设。这个建议不一定总能奏效，但至少当你有意去寻找，而且对目标是什么心中有数的时候，成功的可能性会更高。

线索有时也存在。很多跳出系统的伟大时刻都与抛弃某些备受好评的东西有关，后来我们才发现这些东西其实是不存在的。“燃素”曾被认为是构成火的基本成分，“热质”曾被看作是不可见的、彼此互斥的流体或气体，曾被认为是构成热的主要成分，但这些概念都被放弃了。同样还有“以太”，它被看作一种介质，像声音通过空气和水传播一样，光通过以太传播。其他值得赞赏的跳出系统不是在做减法而是做加法：比如细菌、电子，甚至可能还包括量子力学的多世界诠释！一开始谁也无法认清我们是否应该跳出系统。语言学家雷·杰肯道夫（Ray Jackendoff）和我曾经论证过，我们应该放弃那个人们在讨论意识时几乎总是心照不宣的假设：意识是一种“最高级”“最核心”的心理现象。同时我也论证过：把意识理解为一种以太一般的特殊媒介，经验内容通过这一媒介被转化或翻译，也是一种流毒甚广且未经检验的惯性思维，也应当被打破。像很多人一样，我也曾论证过，如果你认为自由意志与决定论理所当然地不相容的话，那你就犯了一个大错。关于这一点后面我还会再谈。

还有另外一条线索。有时候，一个理论中的问题可能是很久以前埋下的，那时可能有人说：“为了便于论证，让我们假设……”然后人们纷

纷同意：“为了便于论证，我们就这样假设吧。”后来人们你来我往地讨论下去，却没有人记得最初的假设了！我觉得，至少在我们哲学领域，有时候辩论双方是在享受争辩，以至于没人愿意冒险去彻底平息论辩，去检验让整个辩论得以开展的前提。我可以举两个古老的例子（当然它们都是充满争议的）：（1）“为什么存在某些东西而不是无物存在”是一个深刻的问题，需要我们回答；（2）“到底是‘上帝因为某事是善的才命令人做某事’还是‘因为是上帝的命令，所以某事是善的’”这个问题很重要。我想，如果有人能漂亮地回答这两个问题，我再把它们称作“不值得关心的假问题”就于心不安了，可是这不意味着我说错了。谁能说真理必须是有趣的呢？

09

古尔德的3种思考工具
“不如说”“故意堆积”和“古尔德两步”

已故生物学家斯蒂芬·杰伊·古尔德（Stephen Jay Gould）作为“吊杆托架”的开发者和设计师，简直是艺术家级的。在此，为了纪念这位人类中最精通此道的大师，我用“古尔丁”（Goulding）来命名三类有关的思考工具。

1．“不如说”

“不如说”能够轻快地令你滑入“虚假二分”。一般它是这样用的：“情况与其说是如何如何（你所相信的主流观点）；不如说是如此如此（彻底不同于前者）。”有时候这么说确实不错，因为有时你确实不得不二者择一，这时你并没有陷入虚假二分，而是遇到了一个真实的、无法逃避的二分。但有时，“不如说”有点耍花招的意思，因为“不如说”这一短语暗示了：由它衔接的前后两种主张之间有一种重要的不相容性，而这一点无需论证。

我们可以在古尔德对“间断平衡理论”的论证中发现一个标准的“不如说”：

> 与其说变化是以整个物种渐变的方式在不知不觉中发生的，不如说变化发生在孤立的小种群中，从地质学的角度上看，它们瞬间转变成了新的物种。（Gould, 1992b, p. 12）

上面这段话诱使我们相信，进化上的改变不可能既是“在地质学的角度上是瞬间的”又是“在不知不觉中渐变的”。可是这当然是可能的。而且物种的进化恰恰必须如此，除非古尔德想说进化是以突变的方式进行的，在某种“设计空间”中实现巨大的飞跃，可是他又坚持说自己从不支

持“进化突变论”。物种的形成“在地质学的角度上是瞬间的”，意味着它在“很短”的一段时间内发生，我们就假设五万年吧，几乎在大多数地层勘探中都无法探测到这么“短”的时间流逝。在这“短暂的一瞬间”中，我们假设有一种生物的体长从半米变成了一米，它的体长增长了百分之百，但是增长的速率可以是每一世纪增长一毫米，这又让我们不得不说这是一种不知不觉的渐变。

我们再举些其他的例子，争取弄清“不如说”这一伎俩的本质：

> 与其说人类只是一种“潮湿的机器”，就像呆伯特[16]与很多认知科学的研究者所认为的那样，不如说人类拥有自由意志，并且为他们的善举和恶行负有道德责任。

我是还要问，为什么必须二者择一？他并没有论证为什么一个具有自由意志并承担道德责任的人不能也是一种“潮湿的机器”。在这个例子中，作者利用了一个常见的假设，但是这个假设也是有争议的。还有另外一个例子：

> 与其像马克思那样说“宗教是人民的鸦片”，不如说它是一个深刻的安慰性符号，它标志着人类对死亡无可避免的认识。

我还要再问一遍，为什么宗教不能既是“鸦片”又是安慰性的符号呢？我想你们现在已经看到要点了，而且，我们在一篇文献中寻找“不如说”比寻找虚假二分要容易多了，后者可没法这么找：在搜索栏里键入“不如说”并查看搜索结果。需要记住的是：不是所有使用“不如说”的都是我们所说的这种“不如说”；它们之中有些是合理的。同时，有些“不如说”并不使用“不如说”一词，有人会用“是……而不是……”。我可以从某些搞认知科学的空想理论家的著作中摘取一些素材，为大家造一个例子：

> 必须将神经系统看作是主动探测其所处环境，而不是像计算机一样被动地处理由感觉器官输入进来的信号。

谁说处理输入进来的信号的计算机就不能主动进行探测了？被动的、无趣的计算机与主动的、奇妙的有机体之间的常见对立从未得到过恰当的辩护，在我看来这是流毒最广的思维定势之一了。

2．故意堆积

古尔德频繁使用的另外一种思考工具是“不如说”的变种，我称之为“故意堆积”：

> 当我们谈论“从单细胞生物到人类的长途跋涉”时，好像进化是一条沿着不间断的谱系持续进步的道路。没有比这个想法更不现实的了。（Gould, 1989a, p. 14）

没有比哪个想法更不现实的了？乍看起来，好像古尔德说的是在单细胞生物与我们人类之间不存在一条不间断的谱系。可是这种谱系确实存在呀，达尔文的伟大思想所要表达的就是这一点。那么古尔德这里所指的会是什么呢？假设我们把重点放在“进步的道路”上，那么他只是在说相信进步是“非常不现实的”。确实，进化之路是一条连续不断的谱系，但不是一条全球性的进步的谱系。但如果我们不小心谨慎的话，从古尔德的这段话中我们读到的意思就成了：进化论的标准命题有着严重的错误，从单细胞生物到人类之间没有一条连续不断的谱系。但是，我们借用古尔德的话说，“没有比这个想法更不现实的了”。

3．古尔德两步

古尔德还有一个伎俩，被我称为“古尔德两步”，这个思考工具在我以前出版的书里介绍过了。进化理论家罗伯特·特里弗斯（Robert Trivers）在与我的私人通信中给这个思考工具命了名，当然还是以它的发明者命名的：

> 第一步，制造一个靶子，然后驳倒它。这个技巧谁都会。第二步才堪称天才：主动把注意力转移到第一步驳论所用的证据上，这些证据本可以证明你的对手并未持有被你反驳的那种观点，而你则把它们诠释为对手面对你的攻击勉强让步了！（Dennett, 1993, p. 43）

英国达尔文主义哲学家海伦娜·克罗宁（Helena Cronin）写过一本还不错的书《蚂蚁与孔雀》（*The Ant and the Peacock*），它受到了古尔德无礼的批评（Gould, 1992a）。两个月之后，我在给《纽约书评》

（*The New York Review of Books*）的编辑所写的一封信中举了三个“古尔德两步”的例子。在这里我试举其中最简单的一例：

> 最明显的例子是古尔德发明的“外推主义”（extrapolationism），它被说成是对克罗宁的“适应主义”的一个合理的拓展。这种泛延续性、泛渐进主义的教条非常轻易地就被大灭绝的事实所推翻了。“但如果大灭绝确实打断了物种的延续性，如果在正常的时代缓慢建立起来的适应性不能对大灭绝之后的环境做出成功预测的话，那么外推主义将会失败，适应主义也将被打倒。”我无法想象任何一个适应论者会傻到承认任何一种“外推主义”式的观点，就像古尔德描述的那样，这种观念“纯粹”到连大灭绝在对生命之树的修剪中扮演了一个重要角色的可能性都要予以否认。很显然，当彗星撞击地球，发出百倍于我们制造过的氢弹当量总和的爆炸时，即便是进化得再完美的恐龙也必将屈服。而克罗宁的书里没有一个字支持那种错误的观点。如果说古尔德认为进化过程中大灭绝所扮演的角色问题与克罗宁理论的核心问题，即“性选择”理论和“利他主义”有关，他也并没有指出关联何在。在克罗宁那本书的最后一章中，她转向之前并未关注的内容，对物种的起源这一进化论的核心问题做了一番细致的讨论，指出这个问题现在依然悬而未决。但古尔德却像是终于悟到了什么似的，抓住这一章不放，讽刺地说克罗宁在这里承认了她“泛适应主义”的失败。真是可笑！（Dennett, 1993, p. 44）

有一个好玩的项目可以让修辞学的学生来做：把古尔德出过的那么多书都梳理一遍，把他所发明的各式各样的“吊杆托架”编个目录，就从“不如说”“故意堆积”“古尔德两步”开始编吧！

10

小心“当然”这个词
一种让你无需思考就认同的花招

当你在阅读或浏览一篇论证性较强的文章，尤其是哲学类文章时，有一个小窍门可以帮你节省自己的时间和精力，在我们这个使用电脑检索的时代，这个小窍门尤其有效：在文档中搜索“当然”一词，然后仔细检查所有出现这个词的地方。尽管并非次次如此，但“当然”一词通常像盏信号灯一样可以定位出论证当中的弱点所在，它是一个警告标识，提醒人们那里可能有一个吊杆托架。为什么呢？因为它恰好标示出了那种作者确实相信并且期待读者也会相信的东西。如果作者真的认为所有读者都不会怀疑这一点的话，他也就用不着提“当然”了。在此，作者必须决定究竟是对这一点做出解释、提供理由，还是干脆直接做出断言，并理由充分地期待大家同意这一点。毕竟，人生短暂嘛！而恰恰在这种地方，我们会发现一些未经检验的“自明之理”并不是不言自明的！

在评论美国哲学家内德·布洛克（Ned Block）的一篇文章时，我第一次注意到了“当然”是这么好用。在那篇文章中，直接反对我的意识理论的几个重要例子都使用了“当然”一词。为了醒目一些，布洛克还给那段话加了斜体[17]：

> 某些持续的大脑表征足以影响记忆、控制行为，等等。但是，这当然只是一种关于人的生物学事实，而不是文化建构。（*Block, 1994, p. 27*）

这意味着他想要不经论证就把我关于人类意识的理论打发掉，我的理论是人类的意识必须而且实际上就是一系列通过学习才能获得的细微的认知习惯，而不是与生俱来的。我评论道：“凡是布洛克说到‘当然’的地方，我们都得找一找所谓的心智障碍了[18]。”（Dennett, 1994a, p. 549）布洛克是“当然”高手中最肆意滥用的一位哲学家，其他一些人也总依赖它，每当这些人一用“当然”，我们都该敲响小警铃。“这就是无意

识手法发挥作用的地方。轻轻一推，眼睛一眨，错误的前提就被晃过去了。”（Dennett, 2007b, p. 252）

最近，我打算更系统地检验一下我对“当然”的感觉是否正确。我在哲学主题网站philpapers.org的心灵哲学板块中找了大约60篇论文，然后在检索中输入“当然”。我发现大部分文章根本没有用到这个词。我把每篇用了它的文章都检查了1～5遍，大多数明显是无辜的；有一小部分算是有争议；但还有6篇文章，警铃在我听来敲得是又响又脆。当然，对于什么算是明显有问题，每个人都有非常不同的阈值，因此我也没有费心把在这个不正式的实验中的得到的“数据”做成表格。我鼓励善于怀疑的大家展开自己的调查，看看你们能发现什么。在后面我还要仔细分析一个特别过分的例子，详情请见第64章。

11

反问
让你不好意思说“不”

就像应当警惕“当然”一样，你也应当对一个论证、一场争论中出现的反问句提高警惕。为什么呢？因为像“当然”一样，使用反问句是作者想要走捷径的表现。反问句的句尾是个问号，但却不意味着它需要我们去回答。也就是说，作者根本懒得等你给答案，因为答案这么明显，你好意思说不吗？换句话说，大部分反问句都是压缩版的归谬法，明显到不用再说了。所以我们应该培养起一种好习惯：每当看到一个反问句，先试着自己悄悄地找一个不那么显而易见的回答，如果你能找到，就说出来，给对你说反问句的人一个惊喜吧。

记得我很久以前看过一期《花生漫画》（*Peanuts*），它把这一招展现得棒极了。查理·布朗（Charlie Brown）反问式地说道：“那么现在，又有谁能告诉我们什么是对、什么是错呢？”在漫画的下一格中，露西（Lucy）回答：“我呀。”

12

什么是“深马”爱就一个字

我已故的好友，计算机科学家约瑟夫·魏岑鲍姆（Joseph Weizenbaum）一直渴望成为一位哲人，在其职业生涯的后期，他试图让自己的工作从技术性转向富有深意。他曾经跟我讲过一件事，一天晚上，当他在饭桌前眉头紧锁、滔滔不绝地大谈了一番崇高理念之后，他的小女儿米丽娅姆说道：“哇哦！老爸又说了句深马（deepity）！”多么灵机一动的造词啊！(19)我决定采用它，并且把它当作分析的对象。

“深马”，就是一个“看上去”重要、正确、深刻的命题，但它之所以看上去有这种效果是由于它的模棱两可。你从一个角度读这句话，会发现它明显是错误的，但如果它是真的，会显得特别惊天动地；而从另一个角度读这句话，你又会发现它是真的，但是特别无聊。粗心的听众看到了第二种读法中的“真”，又看到了第一种读法中的震撼效果。“哇哦！这句话可深了。”他们想。

这里有一个例子。坐稳了，这句话很给力。

Love is just a word.

爱就一个字。

哇哦！真了不起。很带劲，是不是？错了。在一种读法中，这句话明显是错的。我不确定爱情到底是什么，也许是一种情绪、一种情感依赖、一种人际关系，也可能是人类心灵能达到的最高状态。但不管怎么说我们都知道它不只是一个词。你在词典中可找不到爱情！

我们可以采用另一种读法：用咱们哲学家约定俗成的方式。当我们“谈及”一个词时，就把这个词用引号括起来，那句话就变成了：

"Love" is just a word.

“爱”就一个字。

确实如此，“Love”是一个英文单词，我们可以说它只是一个词，而不是一个句子。它以“l”开头并由四个字母组成，在词典里它一般排在“lousy”与“low-browed”之间，它们也只不过是词。“芝士汉堡”也只不过是一个词。“词”只不过是一个词。

你会说：可是这不公平。显然，无论谁说出“爱情只不过是一个词”时都意指得更多。确实如此，但是他们并没有把这更多的东西明说出来。也许他们的意思是“爱情”是一个像“独角兽”一样的词，它让人们误以为那是多么美好的东西，但事实上却并不存在；也许他们的意思是“爱情”这个词充满歧义，人们永远也无法说清它到底指的是什么样的东西、关系、事件。但这些解释看上去都没有那么有道理。或许“爱情”就是一个令人深感困扰的、难以定义的词，或许“爱情”就是一个难以定义的状态，但很明显，这些解释看起来既没有什么建设性也不深刻。

并非所有“深马”都这么容易分析。最近，理查德·道金斯（Richard Dawkins）[20]又让我见识到了一句“深马”，坎特伯雷大主教罗恩·威廉斯（Rowan Williams）描述自己的信仰时说：

沉默地等待真理，在问号出现之处静坐并呼吸。

我就把对这句话的分析留给大家当作练习吧。

第一部分小结

一件非常趁手的工具几乎可以变成你的一部分，就好像你的手脚一样，对于思考工具来说尤其如此。装备上这些简单的、多用途的思考工具，能让你带着更敏锐的感觉踏上艰难的探索之旅：它们让你看到裂隙、听到警铃、闻到危险的气息，如果少了它们的帮助，你可能无法发现那些错误。

此外，你也要把一些准则牢记于心，比如拉波波特法则、史特金定律，它们会像《木偶奇遇记》中的小蟋蟀一样对着你耳语，当你要挥舞着武器冒失地冲进丛林时提醒你控制自己。是的，思考工具同时也是武器，用战斗来比喻没有什么不恰当的。我们当然需要智性上的雄心与勇气来解决最艰难的问题，而争强好胜就是它们自然的附属品。在激烈的“战场”上，当有人急着要让别人像他一样思考时，即便是伟大的思考者也会采用卑鄙的手段。当人们能够抛出有力的反击时，建设性的批判也会被嘲讽所取代。这样的事我们已经见过太多。

我们接下来要面对的，都是一些热点问题：意义、进化、意识，还有自由意志。当你处理某些境况时可能会感到恐惧或厌恶，不过请你放心，那时你并不孤单。即使是最有名望的专家也可能会被一厢情愿的想法牵着走，因而对真理视而不见，为情感而非理性的原因所蛊惑。人们关心自己是否真的具有自由意志，关心我们的心灵以何种方式存在于身体当中，也关心在这个只由原子、分子、光子、希格斯玻色子构成的世界中意义何以存在，甚至是否存在。人们应该关心这些。说到底，有什么问题会比这些问题更重要呢：我们身处于怎样的世界？在这个世界中我们应该做些什么？所以，我们要格外小心。前路凶险，亦无地图可依。

第二部分

关于意义和内容的思考工具

INTUITION PUMPS

AND OTHER

TOOLS FOR THINKING

为什么要从意义开始说起呢？因为它是所有难题的核心，原因很简单：每当我们开始与自己或者别人谈论它们时，那些难题就会出现。獾不用关心自由意志，海豚也不用烦心意识的问题，因为它们没有提出问题的能力。好奇会害死猫，却能吸引人类去反思自己的种种困惑。现在看来，语言或许会带来一个麻烦：一旦掌握了语言，我们就再也摆脱不了那些大问题了。因为无论怎么看，它们都不再微不足

道。如果可以像我们的祖先类人猿那样注意不到那些问题，或许我们可以过得更舒适，成为更快乐、更健康的哺乳动物。

要进行有效的探索，尽一切所能去弄清出发点和自身装备是我们首先要做的事。词语有自身的意义。那词语的意义是怎样产生的呢？作为词语的使用者，我们通过说某一事物来指示它，而这又是怎样实现的？我们如何理解彼此的话语？我们的爱犬似乎在一定程度上能“理解”一些单词，数量甚至高达几百个之多，但是撇开训练所得的技能和自然界中从灵长类、鸟类到无脊椎类动物的初级符号系统，真正将人类思想与其他动物的思想区分开来的是词语。两者之间差异显著。但有些“较高等”的非人类动物看起来仍像是有思想的，它们也在以一些有限的方式处理意义：它们的知觉状态，它们的欲望和冲动，甚至是它们梦境的意义。

有时，人们会精心打扮，穿上猫装、披上熊皮或者装扮成海豚，那时我们就会觉得自己跟动物很像。所有的人类文明都有这样的共识：动物可以看、知道、想、尝试、害怕、决定、冲动、记忆，等等。总之，它们同我们一样，思想中满是有意义的的某种东西，无论它是观点、信念，还是心理表征。

大脑是如何产生意义的？一个普遍的观点是，既然词语有意义，那么我们和动物头脑中这些代表意义的东西也都能像词语那样，在头脑中组成思想的语句、表达观念，诸如此类。但如果词语的意义来自思想，而词语本身也是思想的产物，那么，这些思想词语又要从哪里来获取意义呢？难道动物的大脑要像词典那样，既保存思想词语，又保存这些词语的释义？还有，既然动物，至少是“较高等”的动物的头脑中已经装满思想词语了，那为什么它们还是不能谈论？[21]思想语言这个观点是非常有问题的，但毕竟我们的思想和信念需要有一些构成要素，它们到底是什么呢？[22]

13

特拉法尔加广场上的谋杀案

意义和内容都是我们大脑的“意向性”产物

这是我们的第一个直觉泵。雅克在特拉法尔加广场上枪杀了他的叔叔，继而被舍洛克逮捕。这则消息汤姆是在《卫报》上读到的，鲍里斯则是在《真理报》上读到的。雅克、舍洛克、汤姆和鲍里斯都有着不同的经历，但有一件事情却是他们所共知的：他们都知道一个法国人在特拉法尔加广场上杀了人。他们都没有说过这句话，即便是“对自己”也没说过；我们还可以猜测，他们的脑中也没有浮现过这一命题，即便浮现过，各自的理解也会不尽相同。而他们确实都知道，一个法国人在特拉法尔加广场上杀了人。

人们从常识心理学（folk psychology）的立场出发会觉得这显然是一条他们共享的信念，但事实上，这种看法是有局限的。普通的常识心理学家，也就是我们所有人，不自觉地就会把这种共识强加于人。以此同时，我们对他人两耳之间生成那些信念的器官却知之甚少。这四个人是不是还有其他什么共有的东西，是否会有某种相同形态的东西在他们的脑中成形，并以某种方式促使他们拥有了共享信念？但这种想法会让我们陷入一种暧昧的推论。也许现实中真的有那种共有的神经结构，四个大脑恰好都以同样的方式“组织”出了这个想法：“一个法国人在特拉法尔加广场上杀了人。”但这不必要，事实上也不太可能。出于种种原因，我们在此只做简单探讨。

“I'm hungry”和“J'ai faim”这两句话使用了不同的单词，大声朗读时发音也不同，各属两类不同的语言，语法结构也有差异，但它们有一个共性：都意味着或者关乎着同一件事——说话者饿了。这一共有的性质，即两句话在各自语言中的意义（语句的意义）或者句子所表达的那些看法（语句的内容）是哲学和认知科学的一个中心论题。这种句子、图画、信念，毫无疑问还有大脑状态所展示的关涉性（aboutness），哲学里称为意向性（intentionality）。其实“意向性”这个术语选择得并不

太合适，因为局外人会习惯性地将“意向”与有意做什么事情这类日常概念混为一谈。

下面这席话可以提示你两者之间的不同：一根香烟不关乎吸烟，也不关乎其他任何事情，尽管事实上它总要被某个人吸掉；而一个“禁止吸烟”的标语就关系到了吸烟，从而显示出了意向性。认为某棵树后面有强盗也显示出了意向性，它涉及了一个可能存在的罪犯，但这显然不是普通意义上的有意为之，你不是“存心那样认为”，只是突然产生了这种感觉。而远离那棵树是普通意义上的有意为之，因为它并不关涉其他什么事情。在遇见哲学术语“意向性”的时候，如果你有用“关涉性”这个稍显笨拙的词替换掉它的习惯，那你就更不容易出错。人们普遍认为意义和内容之间有着密切的联系，它们之间相互依赖，更有甚者还会说，这两者压根儿就是一回事，都是意向性。但对于内容和意义是什么以及怎样准确地捕捉到它们这样的问题，人们却很难达成一致。这就是我们必须谨慎对待这一主题的原因。这里是各路问题的大集会，不过我们尽可以小口小口地细嚼慢咽，一点一点地将它们消化掉。

从上文的例子中我们可以看到，有四个不同的人都知道这起谋杀案的案情，他们的大脑并无共同之处，却共享了同一个“意向性”的内容：相信着“同一件事”。雅克是谋杀案的目击者，事实上，他是犯案人；舍洛克感受这一事件的直接性仅次于雅克；汤姆和鲍里斯则是采用截然不同的方式了解到了这次谋杀。获得“一个法国人在特拉法尔加广场上杀了人”这则信息的方法有无穷多种，利用这则信息以拓展个人能力的不同方法也有无穷多种，比如在智力竞赛中回答与此相关的问题、赢得赌注，或者是将它当作逸事讲给游客们听，等等。纵然我们有很好的理由去相信，所有的信息来源与应用结果都来源于人类大脑的某些共同的结构，而我们最终会发现这些结构，只是，现在就直接跳到这个结论还为时过早。

在结束论述这个直觉泵之前，我们还应该按照侯世达的建议转动那些旋钮，看看直觉泵的各个部分都在做什么。为什么要挑选一个如此特殊的场景？因为我需要的是一件令人难以忘怀的事情，它得足够震撼，以至在远离案发现场的地方也会有关于它的不同语言的报道。需要指出的是，我们的大多数信念并不是这样获得的。也许雅克、舍洛克、汤姆和鲍里斯还有无穷多其他的共享信念，它们看起来并不那么引人注目。例如，他们都相信，椅子比鞋子大，汤是液体的，大象不会飞。如果我告诉你野生的鲑鱼不戴助听器，你可能会说，这并不稀奇。但你是什么

时候知道这件事的？你不可能天生就知道，它也不是学校课程的一部分，你头脑中也从没有出现过表示这个意思的句子。鲍里斯通过阅读《真理报》把那些与事件相关的句子“简单地上传”到了大脑，然后又将所了解的这些东西“转译”成了“大脑语”（Brainish）。对于我们来说，整个过程似乎显而易见。但是，要说鲍里斯的大脑也针对有关鲑鱼的事实进行了这样一番文书工作，就有些不切实际了。这个事实是从什么东西转译成了“大脑语”的呢？

下面是另一个旋钮。假设，一只狗菲多和一只鸽子克莱德也是这起谋杀的目击者，它们可能也从这起事件中接受到了什么，所接受的东西都经它们各自的大脑调整过，会影响到它们以后的行为。但是，即使作用于它们感觉器官的信息有声有色且都指向这一事实，它们大脑最终得出的也不会是“一个法国人在特拉法尔加广场上杀了人”。一段录像可能会为这一事实提供合法的证据，但对菲多和克莱德来说则毫无意义。可见，这个直觉泵是想冒险将一种严重的人类中心主义倾向带入我们对于意义的探索中。词语和句子是典型的意义载体，但动物不使用它们，因此，说“动物的脑会使用语句”即便不算错，也显得有些牵强。而如果事实证明这种说法是对的，那它将会是一个极具启发性的发现，虽然我们之前早已这样认为。

从科学角度来看，意向性现象对我们来说既非常熟悉，就像食物、家具和衣服一般在日常生活中显现，又让你全然难以捉摸。我们毫不费力就能将生日祝福、死亡诅咒和承诺区分开来，但这不代表我们同样可以轻松完成一个工程：做出一个有效的“死亡威胁侦测”。死亡威胁都有哪些共同之处呢？看起来，似乎只有它们的含义。但含义不像放射性或者酸度，很容易就能被调试好的探测器鉴别出来。要检测一种通用意义，我们首先想到的是国际商业机器公司（IBM）的超级计算机沃森（Watson），在按照意义归类方面，它比之前的人工智能系统要好得多。需要注意的是，“按照意义归类”这项工作并不简单，沃森有时可能会错识一些连小孩都能轻易辨别的死亡威胁。一个孩子可能会笑着冲另一个孩子吼道：“如果你再这样，我发誓会杀了你！”小孩子也能识别出这不是一个真的死亡威胁。沃森的庞大和复杂至少间接地说明，“意义”的那些常见的性质其实是多么难以捉摸。

14

生活在克利夫兰的一位兄长
一种观念只能生长在一堆观念之上

还是那个观点，含义并非神秘不可测。我们大脑中一定会有某些结构能够以这样或者那样的方式“贮藏”我们的观念。知道了普度鹿是哺乳动物之后，你的大脑中会有一些东西发生改变；之前一些模糊不清的东西会逐渐变得明朗起来；这些变化还会牵连出一大批关涉物，如此一来才能解释为什么现在你会知道，比起梭鱼类，普度鹿与水牛有更近的亲缘关系。因此人们普遍会认为：观念是“贮藏在大脑中”的，就像数据文件以系统码的形式保存在硬盘里那样，每个个体的系统码也许都如指纹一般各不相同。雅克的观念以“雅克语”的形式刻录在他的脑中，舍洛克的观念也以“舍洛克语”的形式刻录在他的脑中。但是，这个迷倒众人的想法是存在问题的。

现在，假设我们已经进入了神经密码学的繁盛时代。在这里，为观念稍做修补和将观念植入人类的大脑对于“认知微神经外科医生”来说都已成为可能，他们能在人类大脑的神经元中写下相关的命题，当然，用的是大脑中的“本土语言”。我想，但凡我们能够读懂大脑中的文字，那么，只要工具足够精妙，我们也一定能书写它们。

来做个假设，我们要把“我有个兄长生活在克利夫兰”这个错误的观念植入到汤姆的大脑中。假设认知微神经外科医生已经按要求非常完美地完成了这次改写。之后可能会出现两种结果：这种重新改写已经损害到汤姆的基础理性，又或者没有。请考虑这两种情况。汤姆正在一家酒吧里坐着，一个朋友问：“你有什么兄弟姐妹吗？”汤姆回答说：“有，我有个哥哥住在克利夫兰。”“他叫什么名字？”那么接下来会发生什么呢？汤姆也许会说：“名字？谁的名字？哦天哪，我说了什么？我根本就没有什么哥哥！刚刚，在那里，我恍惚觉得我好像有个哥哥住在克利夫兰！”在另一种情况中，他可能会说：“我不知道他的名字。”逼得紧了，他甚至会全盘否定这个兄弟，或者这样断言：“我是独子，我还有个哥哥

住在克利夫兰。”

在这两种情况中，我们的认知微神经外科医生都没有让一个新的观念很好地奏效。在第一种情况中，这个单独的、无法证实的干扰一出现就被汤姆完整的理性彻底粉碎。短暂情形下说出的“我有个哥哥生活在克利夫兰”不是一个真正的信念，它更像是一种抽搐，像是抽动秽语综合征的一种表现。而在第二种情况中，如果可怜的汤姆继续表现出这种病理特征，那么，在有关哥哥这个命题上所表现出的赤裸裸的非理性会让汤姆无法再成为一个可以持有观念的人。无论是谁，如果理解不了“有一个哥哥住在克利夫兰自己就不可能再是独子”，那他也就不能理解自己所下的断言。没有理解，你就只能“鹦鹉学舌”，无法形成观念。

这个假想出来的例子凸显出了观念之下隐含的一种心智能力。如果一个观念不能根据不同的语境无限扩充自身的使用方法，那么，在任何意义上它都不能算作是一个观念。假设那位医生已经完美地植入了“我有个哥哥生活在克利夫兰”这个观念，同时也保留了大脑的正常功能，在这种情况下，一旦出现了什么矛盾，大脑便会立即撤销掉这个人工植入的观念，否则，它将会表现出一些病理特征，围绕这个植入的观念展开层层叠叠的虚构。比如：“他叫塞巴斯蒂安，住在热气球里，是马戏团的杂技演员。”这样的谈话并不少见。“科萨科夫综合征”是常常困扰嗜酒者的一种健忘症，患者都深信自己用虚假“记忆”编织起来的那些故事。

所有这些都清楚地说明了，一个“命题”不可能孤立地存在于我们的大脑中，即使是一个妄想出的观念，它也只有在一堆非妄想的观念上才能生长，它得承接住那些不断传递过来的意义。如果在汤姆不能同时确信自己的哥哥是男性、有呼吸、住在波士顿以西巴拿马以北等的情况下，就说医生成功植入了一个观念，那可就大错特错了。

这一直觉泵想要说明的是，人们不可能只持有一个观念。你知道狗有四条腿，就一定知道腿就是狗的四肢，四一定比三大，等等[23]。每个观念都会牵涉到其他一些东西，这里我就不再一一列举了。在此，我也不再赘述要怎样改造这件思考工具，让它发挥其他作用。欢迎你自己转动旋钮，看看还会想到些什么。在更多地描述出它们的特性之前，我要尽可能多地先把各种各样不同的思考工具展示出来。

15

“爸爸是名医生”

理解是随着时间推移而逐渐清晰的

问一个小孩：“你爸爸是做什么的？”她回答：“爸爸是一名医生。”可是她真的知道她在说什么吗？从某种意义上看，她当然知道，但要想真正弄懂自己所说的这句话，她还需要知道些什么呢？同理，如果她回答的是“爸爸是一名套汇人”或者“爸爸是一名精算师”，她还需要知道些什么？假设，我们怀疑小孩可能并不明白那句话的意思，想测试一下她，难道她就一定得对那句话做出释义，或者必须要说出她的爸爸给病人治病，来对之前的回答做出扩展吗？如果她是因为知道爸爸不是屠夫、不是面包师、不是烛台匠，从而得出爸爸是名医生，这样的理由充分吗？如果她不知道假医生、江湖医生、无证医生这些概念，她会知道医生是什么吗？

就此而言，要搞清楚爸爸就是自己的父亲，她又需要理解多少事情？要搞清楚他是自己的养父，或是自己“生物学上的”父亲呢？显然，对于成为一名医生意味着什么，作为爸爸又意味着什么，小孩子的理解会随着时间的推移逐渐深化，所以，对于自己的这句“我的爸爸是一名医生”，她的理解也会有所变化。我们是否能非常审慎地详细列举出，要“彻底地”理解这一命题，她到底要知道多少东西？

如这个例子所示，因为理解是一点一点形成的，那么基于理解的观念也应该是一点一点形成的，即使是对于这种通俗的命题来说也是如此。她“近似”知道她的爸爸是一名医生，这并不是说她对此有所保留或者怀疑，只是她的理解还不够全面，还不能作为一个重要的前提引出其他有用的观念。

16

常识映像和科学映像
反映世界的两个不同视角

到了搭建一些脚手架的时候了，然后再继续我们的探索，这样可以更好地理解什么是意义。这里有一件思考工具，它能为许多问题提供有价值的研究视角，可惜，本应人人必备的这样一件思考工具至今却只局限在了哲学的本土领域中使用。哲学家威尔弗里德·塞拉斯（Wilfrid Sellars）于1962年发明了它，旨在能够更清楚地考量：对于我们居住的这个世界，科学向我们展示了什么？常识映像是我们日常接触的那个世界，它包括固体事物，还有颜色、气味、味道、声音、阴影、植物和动物，也包括人和他们所需要的东西：包括桌子和椅子，桥梁和教堂，金钱和契约，还有歌曲、诗篇、机遇或者自由意志这些无形的东西。

然而，当我们将所有这些事物与科学映像中的那些东西，如分子、原子、电子还有夸克及其家族对应起来时，会出现哪些疑难问题呢？其一，真的有什么东西是实在的固体吗？20世纪初，物理学家亚瑟·爱丁顿爵士（Sir Arthur Eddington）曾写道：有“两张桌子”，一张是我们每天都触摸的固体的那张，另一张由原子构成，这些原子散布在一片空旷的空间之中，它们组成的与其说是一块木料，倒更像是一个银河系。有些人会说，科学想要说明的是，世上没有真正的固体，固体性质只是一种错觉。但爱丁顿没有走得那么远。其二，有些人会说色彩是一种错觉，真是那样吗？人类视觉确实只能看到介于红外线与紫外线之间的狭窄波段的电磁波，这种电磁波不含有一点儿带颜色的东西，原子也没有颜色，即使金原子也不是金色的。但从另一种意义上看，颜色当然不是一种幻觉：当索尼公司（Sony）说他们生产的彩色电视能够真实地显示出世界的颜色时，我们不认为这是在说谎；也没有人会因为宣伟公司（Sherwin-Williams）向我们兜售的各异彩色只是些涂料就以诈骗罪把它告上法庭。

那么钱币呢？目前，没有多少钱币是由银制成的，有的甚至连纸币

都算不上。就像诗歌和诺言那样，它们不是物质，承载的是信息。但这就意味着它们只是虚幻吗？不，如果你因此就执意去寻找构成它们的分子，那你就错了。

塞拉斯（Sellars, 1962, p.1）有句名言："理论上来说，哲学的目的就是理解最广泛意义上的事物是如何在最广泛意义上勾连起来的。"这是我读到的最好的哲学定义。科学家们并不特别擅长将我们常识映像中熟悉的事物转换成相对陌生的一些科学映像。物理学博士，请你告诉我颜色是个什么东西吧。根据你的理论，真的存在颜色吗？化学博士，你能为便宜货给出一个化学公式吗？便宜货当然（叮！）是存在的。那它们是由什么构成的呢？嗯，也许根本没有什么便宜货可言。但是这样一来，在一个便宜货和一个看起来像便宜货的东西之间有什么化学成分上的区别呢？

我们可以以这种方式一直追问下去，继续探索那些只有哲学家才会考虑去解决的问题，但现在还是让我们遵循塞拉斯的引导，暂时退后一步，去了解"这个世界上有两种完全不同的视角"吧。那么，为什么存在两种视角？或者，会存在多种视角吗？我打算先从科学映像入手，然后顺势摸索常识映像，并通过这种路径来尝试回答这些问题。

不管是细菌还是我们现代人中的一员，每个生物体在这个世界中都有一套对自己至关重要的东西，对于这些东西他们必须竭尽全力去辨识和抢占。哲学家们将支撑起存在的这一系列东西叫作"本体论"（ontology，源于希腊文的"事物"一词，想不到吧）。所以，每一种生物都有自己的本体论。它还有另外一个名字，叫作生物体的"环世界"（Umwelt，德语单词，意思是"环境"，想不到吧；Thure von Uexküll，1957）。动物的环世界首先是由一些基础情景构成的（Gibson, 1979），包括食物、配偶，它们要学会躲避，知道从哪里钻进钻出，从哪里观测外界的动静，要有一个容身之所，还要有一席立足之地，等等。

从某种意义上看，一个生物的环世界其实是一种内部环境，是一种"主观"甚至"自恋"的本体论，构成这种本体论的都是生物认为对自身最有用的那些东西，但这种"内在"和"主观"并非是从"可意识到"这个层面来说的。环世界其实是一个工程学概念：一部由电脑操控的升降机的本体论，是指在完成这项工作时需要它逐个完成的那一整套动作。[24]图勒·冯·于克斯屈尔（Thure von Uexküll）有一个研究就是关于虱子的环世界的。我们可能会认为，与我们人类的环世界比起来，海星、蠕虫或

雏菊的环世界与那部升降机的本体论更接近，但事实上，这只是我们的常识映像。

与雏菊的本体论或者环世界不一样，我们的常识映像是明显的，也的的确确是主观的。世界是我们所在的世界，世界取决于我们。[25]与雏菊的本体论相同，我们的常识映像也是在经年累月的自然选择中成形的，并最终沉淀成了我们基因的一部分。有一个我很喜欢的例子，讲的是一只食蚁兽和一只食虫鸟的环世界到底有多不同（Wimsatt, 1980）。一只在空中搜寻单个虫子的鸟，为了追踪虫子飘忽不定的飞行轨迹，它的视觉就需要有很高的闪光融合率。鸟儿每秒可以看到比我们的肉眼所能捕捉到的更多的帧数，所以一部电影对它们来说就像是幻灯片一样一片片地切换展示。食蚁兽只需要分清蚂蚁出没的大致区域，然后用自己巨大的舌头一扫蚂蚁密集的地方。

哲学家们会说，在食蚁兽看来，“蚂蚁”是一个集合名词，就像“水”“冰”“家具”这样的词，而不同于“橄榄”“雨点”和“椅子”这类可以一个一个数清楚的词。食蚁兽在看到一大堆蚂蚁的时候会不假思索地用自己的舌头将它们一扫而光，根本不在意什么个体，正如我们在吃糖的时候也绝不会想，自己嘴里咀嚼的是一大堆糖分子。

我们的大多数常识映像并不能随基因遗传，它们基本上是我们在童年时代经过反复灌输保留下来的。词语对我们来说非常重要，有它们作为媒介，我们大量的常识映像才能够传达。但是，我们用词语区分世界中不同事物的本能、我们对说话的冲动，这些多少还是跟基因遗传有些关系的，就像鸟儿能定位飞在空中的一只只小虫，黄蜂总是要筑起蜂巢一样。面对妈咪、小狗、饼干，即便没有语法将它们编组成句子，这几个词仅作为标签，也能帮我们在常识映像中清楚地归拢出很多重要的范畴。如果不借助这几个词本身，你能清晰地搭建出便宜货、错误、许诺的意思吗？更不要说全垒打或者单身汉这类词了。之前通过仔细考量侯世达那些心爱的思考工具，我们已经展示出：词语怎样建构和修饰了我们的思想，怎样运用那些几乎无形的东西，那些天马行空、空口无凭、相互影响的东西，来丰富我们个人的常识映像。

17

常识心理

人们无需正规教育就拥有的一种能力

常识映像中最重要的一种形式是常识心理（folk psychology），因为它还维系着很多其他类别的映像，而这些映像对我们来说也都非常重要。1981年，我提出了常识心理这一术语，但很显然，它的雏形Volkpsychologie早在心理学家威廉·冯特（Wilhelm Wundt）和弗洛伊德等人的作品中就出现过，表示的是有关国民性格的一些东西。我当时没太注意这个词，而也有些人将它沿用至今。我用常识心理这个词表示我们共有的一种能力，我们通过这种能力来理解周围的人、动物、机器人，甚至是一个普通的温控器，认为他们都是拥有关于自身所在世界的知识（观念）、具备努力要去实现目标（愿望）的行动者，会根据自己的观念和愿望选择最合理的行为方式。

一些研究者喜欢将常识心理称作“心智理论”（theory of mind，TOM）。我觉得这极具误导性：对于我们如何才能拥有这种能力的问题，“心智理论”预设了我们运用的是一种理论。这就相当于，如果你会骑自行车，就说明你已经有了一套自行车理论，或者说，你没有饿死或者没有吃沙子，是由于你有一套营养学理论。对我来讲，要思考这些能力，这显然不是一个有效的方法。每个人都相信我们拥有相互理解的能力，但对于我们要怎样才能做到相互理解，大家却众说不一。所以，最好还是不要出现“理论”这个词，于是我选取了“常识心理”这个更加中性的词语。

同样研究如何诠释和预测他人心灵的理论还有心理学，也叫作科学心理学，它确实有自己的一些理论，如行为主义、认知论、神经计算模型、格式塔心理学，等等。而常识心理是一种不需要接受正规教育就可以熟练掌握的能力。常识物理（folk physics）亦是如此，所以拥有这种能力的我们能预料到液体会流动、物体无支撑会坠落、炙热的东西会烫伤我们、水能解渴、滚动的石头上不长苔藓等。至于为什么无需参加物

理课程，我们的大脑就能轻而易举地做出这样几近完美的预测，这又是另一个非常有趣的问题。

常识心理是指“每个人都知道的”那些我们自己及别人心中都有的想法：人们能够感觉到并分辨出疼、饿、渴这些不同的感觉，能够记住过去发生的很多事件，也能预见到一些事情，睁开眼能够看见眼前的事物，在听力所及的范围内能听到别人在说什么，他们骗人也被人骗，知道自己在哪儿，能认出别人……我们不知道其他人、其他动物的脑子里在想什么，就以惊人的自信做出了这些推测。这些本事理所应当地发挥，以至于能注意到它们都是件让人觉得很费劲的事儿。

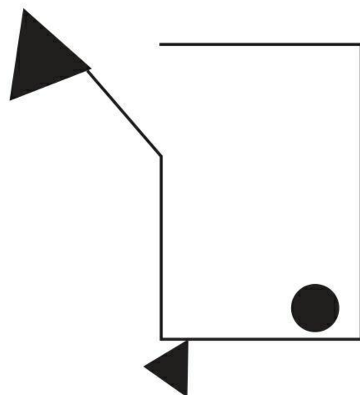
艺术家和哲学家们一致认为：“将熟悉的变陌生”[26]是他们必须主动完成的一项任务。创作天才们的一些奇思妙想威力无穷，能让我们冲破过度熟悉造成的坚壳，跳出系统，转换到一种新的观察视角，用全新的眼光再去看那些平凡无奇、显而易见的事物。科学家们也完全认同这一点。牛顿就问过自己这样一个奇怪的问题：为什么苹果会从树上掉下来？而我们普通人只会想：“那么，怎么就不会从树上掉下来呢？因为它有重量呀！”好像这样解释就足够了。如果你不是盲人却有几个盲人朋友，那么你免不了会像我说的这样：即使眼前的人看不到，你也禁不住会用手去指东西或是比划着各种形状，不管你们彼此相伴了多久。在与一个意识清醒的人共处时，你就是会有这种默认的预期：你们能看到同样的事物，听到同样的声音，闻到同样的气味。

你驾车以60千米每小时的速度在道路上行驶，而另一辆车以相同的速度迎面向你驶来，你不会因此就惊慌失措。问题来了，你怎么就知道不会发生可怕的相撞呢？因为无需思索你就会这样想，虽然你肯定不认识对面的司机，恐怕都无法看清这个人，但他肯定想活下来，而且你俩都知道此时最好的方法就是双方都靠右行驶。值得注意的是，听到广播新闻说今天在这条路上会有新型机器人的驾驶测试倒有可能会让你心情紧张。当然，进行这种系统测试是想让路面交通变得更加安全，毕竟这是谷歌（Google）的设计。但是，与众人鼓吹的机器人比起来，你更相信的是大家在一般情况下的合理行为和自己对迎面而来的这位驾驶员的基本认识。

为什么“每个人都知道”常识心理呢？它是否终究还是一种大家在孩童时代都学习过的心智理论呢？它是我们天生就知道的吗？或者，如果是习得的，我们是什么时候，又是怎样学会它的呢？在近30年中，有大量的研究在探讨这些问题。

我的答案经过了多年的研究，始终未变：与其说这是一种理论，倒不如说它是一种实践，常识心理是我们探查世界的一种方式，它来得那么自然，所以在我们的大脑中一定有着其基因基础。还在妈妈膝边玩耍的时候你就会接触到一些常识心理，之后，如果在成长期间脱离了与周围人的联系，那么你在常识心理方面可能就会表现得比较吃力，在其他方面也会有严重的障碍，但是，你还是会不自觉地把那些看起来不规则运动的事物解释成行动者，并与钟摆或是滚下山的石头区分开来。

1944年，弗里茨·海德（Fritz Heider）和玛丽-安·西梅尔（Mary-Ann Simmel）在心理学入门课程上展示的那段简短动画也很好地说明了这一点：动画中有两个三角形和一个圆形，它们围着一个开口的长方形不停地绕进绕出。这些几何图形怎么看都不像是人，也不像动物，但我们不可避免地会将它们的交互运动看成是有目的的，受欲望、恐惧、勇气或者愤怒驱使而做出的行为。



我们生来都具备一部“行动者检测设备”（Barrett, 2000; Dennett, 1983），它一触即发。但这种设备有时也会走火，这种情况常常发生在一些令人感到紧张的环境中，于是我们就会看到幽灵、地妖、小精灵、矮妖、花仙子、小矮人、恶魔以及诸如此类的事物，而事实上，我们在那个地方看到的也许就只是几根摇动的树枝、一排倾倒的墙，或是一扇吱呀作响的门而已（Dennett, 2006a）。从很小的年纪起，我们就能毫不费力地自动将他人看作行动者，而且不仅能看出他们是高兴、愤怒、困惑还是害怕，也能看出他们在偷偷摸摸地行动、犹豫该走哪条路，乃至不愿接受一笔交易。不用接受脑部手术，也无需多么复杂的技术，这很简单。我已经说过，是那些简单化的假设造就了常识心理，让它们容易使用又力量强悍。它们就像是科学中那些理想化的模型，极度抽象到只剩下了一些本质特征。我称它为“意向立场”（intentional stance）。

18

意向立场
解释实体行为的一种策略

到目前为止，一切顺利。基本上，我们每个人对常识心理都很在行[27]。在大多数时间里，我们都有本事把包括自己在内的所有人看作是有思想的人，这如同呼吸一般毫不费力。我们依赖这个想法，无需思索，也没有疑虑，而通常情况下，它也的确可靠。为什么我们自然而然就会这样想？这种想法又是怎样运行的呢？现在，我们需要停下来多建几个脚手架，好让我们接下来的探究轻松些。

常识心理是怎样运行的呢？我姑且先不让它运作在我们人类身上，检查它在物体上的运作更方便我们得出一个清晰的答案。假设你正在与一台计算机下棋。你想要赢，而战胜对手的唯一方法是你能够料想到这台计算机可能回应你的每一步："一旦我把象移到那里，它就会被计算机吃掉；而如果我移动卒，计算机就必须得动它的后；……"你怎么知道计算机会那样走呢？难道你朝它里面看过？你曾研究过它的下棋程序？当然没有。你不需要那样做。能如此肯定地做出这些预示是因为你是这样设想这台计算机的：

1．它"知道"游戏规则并且"知道"怎样下象棋。

2．它"想"赢。

3．它能够"看到"可能的棋步和机会，并以此来判定如何理性地走下一步。

换句话说，你已经把这台计算机设想成是一名优秀的国际象棋选手，至少不是国际象棋白痴或是那种自杀型的选手。也就是说，你把它当成了一个有思想的人。再往前推进一步，一旦你开始用常识心理来预测和理解这台计算机的每一步，你就已经采用了意向立场。

意向立场是一种解释包括人、动物、人工制品等实体的行为的策

略，它将这些实体的行为解释成一个理性的主体在“考量”了自己的“信念”和“欲望”之后所做出的“行为”“选择”。(28)给这几个词语加上了引号是想强调，我们需要先将它们的标准内涵搁置一旁，以便探索它们的中心要素：它们在实际推理中的作用，以及之后它们对推理者实际行为的预示。按照定义，任何事物，只要从意向立场出发具有可靠的、强大的预见性，那它本身就可以构成一个意向系统（intentional system）。我们将会看到，有很多令人着迷的、结构复杂的东西，虽然它们没有大脑、眼睛、耳朵和手，所以根本不具有思想，但也能自成一个意向系统。也就是说，在人类互动交流以外的其他一些领域，常识心理也可以发挥它的基本作用。我们不仅会看到它在计算机技术和认识神经科学领域里的应用，还会看到它在进化和发展生物学中也发挥着作用，这是常识心理得到充分应用的最重要的几个领域。

我计划先简单搁置起“谁真正地拥有思想”这个麻烦的问题，实际上，它与意向立场适用于哪些领域有关系。但不管这个问题的正确答案是什么，也不管它是否真的有正确答案，都不会动摇这一朴素的事实：除了能在我们日常的人际生活中很好地发挥常识心理的作用以外，意向立场在其他一些领域也一样发挥着十分显著的作用。

搁置问题这一举动让很多哲学家懊恼至极。他们想要阻止我，在进行下一步的探讨之前，他们坚持要先就“什么是思想，什么是信念，什么是欲望”这些问题给出恰当的解释。先生，请先定义一下你的这些术语！不，我不想这么做，这还太早。意向立场是个不错的技巧，我想先探讨它的威力和应用范围。一旦搞清楚了它的长处以及为什么它会有这样的力量，我们就可以回过头重新考量，我们是否还需要对那些词语给出正式的、无懈可击的定义。

先把“谁真正地拥有思想”这个问题搁置起来，面对一个棘手的问题，我们得一点一点地攻克，我不想费力地将它囫囵吞下去、一口消化掉。在我要介绍的思考工具中，很多都可以用来这样一点一点地解决问题，你可以先粗略地“确立”几个点，方便认清问题的大体轮廓。在《活动空间》（*Elbow Room*，1984a）一书中，我将我的方法与雕刻家的方法进行了对照，他们先在大理石上勾画出要雕刻的形象，然后再仔细地、认真地雕琢，慢慢成型。这是一个循序渐进的过程。很多哲学家好像不会使用这种工作方法，他们一定要为自己的问题或者可能的解决方案找定一个完整的界限，然后才肯提出假设。

三种立场

现在我们要比较一下意向立场与其他预见策略，看看意向立场的威力到底来自哪里。首先要做的是辨别三种主要的立场：物理立场、设计立场和意向立场。它们还可以根据不同的情况做进一步地细分，但这里暂不需要。

简单来说，物理立场就是物理科学的一种标准的运算模式，它要求我们在思考问题时用物理定律和事物的物理法则来得出预测。你手一松，石头就会掉到地上，我说这话的时候动用的就是物理立场。一般说来，面对着那些无生命的事物或是非人造的事物，只能采用物理立场，尽管我们在后面还会看到一些很有分量的特殊情况。而对于所有的物质，不管它是不是人造的、是否有生命，因为它们都符合物理法则，所以，理论上，只要从物理立场出发，我们就能够对它们的运动进行解释和预测。如果你手中攥着的是一个闹钟或是一条金鱼，我会基于同样的物理立场得出相同的预测：一松手它们就会掉到地上。从物理立场来看，你手里的这个闹钟或是这条金鱼很少会有其他有趣的反应方式。

闹钟作为一种人造物品，与石头不同，它也服从于人们设想出的预见。这种预见来自于设计立场。假设有一件我从没见过的东西，大家称它为闹钟，那么我一定很快就能推断出：如果我按照要求按动一些按钮，一段时间后这个闹钟就会叮铃作响。我不需要学习那些专门的物理定律去解释为什么我上了闹铃闹钟就会按时响起。我只需要知道，它是一个特殊的装置，会依照相应的设计按要求工作。从设计立场出发的预想没有从物理立场出发的预想那么可靠，原因是我们还需要考虑如下的这些额外的设定：

1．这个实体是按照我们的想法设计出来的。

2．它会按照设计要求运行，中间不会出现失误。

设计出的物品偶尔会出现设计上的失误，有时它会突然失灵。这些丧失功能的情况不会发生在一块石头上，因为石头自身不带什么功能设置，一块石头一旦损坏，结果就是它分成了两半，变成了两块石头，而不是一块损坏了的石头。面对一件相当复杂的设计品，比如说，比一把斧头更复杂的一把链锯，此时极为简便的预期就补偿了采用设计立场所冒的风险。如果我们参考一张简单的设计图纸就可以弄明白链锯的各个

部件如何活动，那么就没人愿意从那些基础的物理定律入手再去预测链锯的运作方式了。

使用我们讨论的重点，意向立场比使用设计立场要承担更大的风险，但也更加便捷。意向立场是设计立场的一种，意向立场将设计而成的事物看作是行动者，有信念也有欲望，能根据这些信念和欲望执行合理的行为。闹钟是一种简单的物品，严格说来，在理解为什么它是那样运行的时候，我们没必要运用奇幻的拟人理论。不过，一旦问题中所涉及的人工制品比一只闹钟要复杂得多时，采用拟人理论这样一种蓄意立场就变得非常必要了。事实上，那几乎是必须的。现在，再借用与计算机下象棋这个例子，让我们来慢慢梳理一下，不要错过任何重要的细节：

① 在轮到计算机走棋时，你要将计算机可用的有效走法一一列出，一般来说会有几十种走法。

② 将这些走法从好（最明智、最理智的）到次（最愚蠢、自杀性的）依次排列。

③ 做出你的预测：计算机会选择最好的那种走法。

也许你一时确定不了为计算机列出的这些有效走法中哪种是最好的，计算机倒是可能会更清楚地“领会”到自己的处境，但你却可以排除掉大多数的可能，只留下四五种走法，这就大大提高了你正确预测的概率。在这个基础上继续思索，借助于设计立场，你就能预先确定出计算机接下来的走法，尽管这可能需要不断的努力、耗费大量的时间。你需要先获取程序的“源代码”（见第27章），然后对其进行“手动模拟”，经过数百万个甚至上亿个小步骤，最终确定出要走哪一步。这样你就一定能预测出计算机的走法，但时间有限，在有生之年恐怕你是不能像计算机那样得出结论的——信息量太大！但相比起要回落到物理立场，奋力处理那些由计算机的键盘输入所形成的电子流，运用设计立场已经让问题简单了很多。所以，试图运用物理立场去预测或解释那台计算机的下棋招式显然不切实际，即使是从设计立场出发工作量也太大，我们需要也有一台计算机，但那是犯规的。

意向立场则能巧妙地完成所有那些繁重的信息收集和运算，将赌注押在一件事上：这台计算机足够“理性”，能够走出最漂亮的一步棋。计算机的目的也是要赢，而且它也熟悉棋盘上每一颗棋子的地位和作用。

在很多情况下，利用意向立场你不用费太多力气就能非常准确地预测出计算机会如何走棋，更何况，有些时候计算机的布棋走势会很明显，是一步“迫棋”，或者叫“无脑棋”。

很明显，在与电脑下棋的过程当中，从意向立场出发可以非常有效地对计算机的走棋进行预测，因为这个下棋软件追求的只是在一个纯理性的下棋环境中“推理出”最合理的那一步。假如计算机程序操纵的是一家炼油厂，同样明显的是，它得根据对各种条件的探测来执行不同的步骤，这些条件多少会影响到程序下一步的操作。这里需要强调的是设计的卓越性或合理性，一个白痴程序设计员设计出来的程序往往很难在这种环境下完成专家们要求的那些事务。而如果是一套设计精良的信息系统或者控制系统，那么即便编写这个程序的程序员们没有为源代码附上一些“注释”说明，程序运行的那些基本原理也都会清晰可见，且具备高度的预见性（后面我们还有很多有关计算机的话题）。要预测程序的运行结果，我们无需清楚计算机的编程，只要知道运营一个炼油厂所需要达成的理性要求即可。

现在大家已经看到，为什么我们可以通过意向立场来预测行为，以及我们是怎样做到的。在我们把人类都处理为意向系统时，我们会巧妙地将对方脑壳中那颗大脑以及我们自己大脑中的运作细节忽略掉，因为我们都在无限制地依赖“人是有理性的”这一事实，浑然而不自知。所以，就算突然闯入了一个新异的情景，我们通常也能自然地，事实上是不由自主地寻求到其中的意义，凭借的正是我们天生的本领：领会别人应有的信念（摆在他们眼前的事实）和欲望（什么对他们有好处）。

没有多少人会质疑“我们常常需要利用常识心理来进行预测”这件事情，至于要怎样来解释这份生命的馈赠，其间却存在着很多争执。“自然法则”成千上万，比如：

> “如果一个人很清醒地睁着双眼，面对着一辆公共汽车，那么他就能知道自己面前有一辆公共汽车。”
>
> “如果一个人发现与别人合作不用花费太大代价，还能得到他人青睐，就会倾向于合作，甚至是同陌生人合作。”

我们是这样一条条学到它们的吗？还是，基于“人们会在情境中根据理性做出回应”这种隐含的观念，所有那些基础法则才应运而生？我保证是后面一种。原则上，我们会在生命历程中逐一领略到概括性的法则，

这些法则中囊括大量定型的行为模式，从中很难产生与人类正常的生活处境相差极大的、科幻小说式的情节，让我们无法想象身处其中时会怎样反应。“如果这事发生在你身上你会怎么办”是一个很平常的问题，有时我们会做出这种无用的回应：“我可能会昏死过去了。”有时则会有效地做出理性的回答：“希望我足够聪明，知道该如何去做。”有时，我们会在一些反常规的场景中看到各色人物创造性的表现，但我们仍然能够顺利地理解他们在做什么、为什么那么做。我们也能理解那些用自然语言表达的全新句子，而这些句子我们在之前的生活中并没有听到过，但我们依然能理解那么多的人类交流，这些现象足够表明，每个人多多少少都与生俱来地具备一些生成能力。

我们很自然地、不假思索地就会将意向立场蔓延到动物那里去。如果要设法捕获一只狡猾的野兽，利用意向立场对我们来说就是不二之选；如果要想理解一些更简单的动物乃至植物的行为，采用意向立场也是一个有用的方法。蛤蜊有自己的行为方式，它们是理性的，这些理性基于它们对这个世界的有限视野。树木能够感觉到敌人在步步逼近，现在洒在自己身上的阳光就要被身旁那些更高大的绿植反射掉了，于是它们吸收养分更快地长高，这并不稀奇。毕竟，对一棵植物来说这是眼下最明智的做法。就人工制品来说，即使原始的温度调节器也会表现出一点最基本的意向立场来：为了要把温度控制在我们想要的水平上，它会对实际温度进行有规律的采样，将采集到的数据与我们的要求进行对比，并依据结果进行温度调节。在给没有专业知识的小朋友介绍温度调节器的工作原理时，你可能就会这么说。

这个简单的意向系统理论讨论的是，为什么将对象视为行动者，我们就能够理解它的行为，我们是怎样做到的。这不是一个关于内在机制的理论，不能直接解释行动者如何以某种方式获得理性指导从而显示出预见性。意向立场向我们提供的是意向系统的一纸“说明”、一份工作说明书，包括哪些需要甄别、哪些得记住、哪些要去做，等等，要实现说明书中所述功能的是工程师，如果是一套有机体意向系统，则是进化和发展过程。眼前有这么一个行动者，它能区别一美元和十美元，可以找零，还可以识别假币，它愿意也有能力一天24小时地为消费者提供他们想要的产品。这种从意向立场出发的特征描述，既可以看作是一份自动贩卖机的说明，又可以看作是对一名便利店店员基本工作的描述，它完全模糊掉了个体之间内部结构的不同和他们具有的其他更多的能力。

用软件工程师的话说，这种等同或者说中立性不是一种错误，而是

一种特征。在后面三章中我们将会看到，正是这种等同让意向系统理论在有关思想和大脑的困惑之间架起了桥梁。简而言之，它让我们看出了“真正的”信念（人的信念）和“仅仅”像是有信念（比如自动贩卖机、动物、小朋友，最明显的还有构成人体的各个部分）两者之间的共同之处。为了进行预测，我们可以借助于意向立场，先简要介绍大脑中那些子系统的功能，而后再对如何运行这些功能给出详尽解释。我们把一整个神奇的、“真实的”人分解成一些仍带有自身程序和方法的次人类行动者，然后再依次将这些次人类行动者分解成更简单、更愚钝的行动者，最后剩下最简单的意向系统，我们无需借助意向立场就可以对它进行描述。受益于有原则地放宽了那些哲学家们曾力图施加在人类信念和欲望上的条件，想象力终于得以在人类个体层面的常识心理和次人类层面的神经回路活动之间架起桥梁，这项任务错综复杂。从人类一路向下，直到无感觉的物体，“真正的”信念和欲望在何处终止，仅仅“好像”信念和欲望的东西从何处开始出现呢？正如我们已经在第15章以及即将在其他一些思考工具的帮助下看到的，要求得出这样一条明确的界线并不合适。

意向立场在计算机科学和动物心理学方面都有非常普遍的应用，意向系统理论解释了其中的缘由。有些进化生物学理论家扬言要摒除意向立场，这真是自寻烦恼，我们在后面讨论进化论的部分会看到这一点。[29]

19

人与“次人”的区别
自上而下的认知能力分解

不是你的眼睛看见了，而是你看见了；不是你的嘴在品尝巧克力蛋糕，而是你在品尝；你的大脑不会憎恶你肩膀上那穿刺般的疼痛，而你会；签订那份合同的不是你的手，而是你；你的身体只是被唤醒，而你却堕入了爱河。前面的句子不只是一种“语法上”的修辞，就像在有雷阵雨时我们会说“天下雨了”而不是“一场雷暴雨正在下”。这关乎的也并不只是一些约定俗成的使用习惯。有时，人们会很不屑地问：“这不就是些语义学的东西吗？”他们倾向于认为，这些与我们如何“定义我们的词语”无关。但是，我们如何定义词语却常常能够产生重要影响，接下来我们就会看到这一点。我们怎样谈论别人，怎样谈论他们能做什么、会有什么样的遭遇，这些都要在基于很多重要的事实之上的前提下才能实现。

乍一看，有很多人能做的事情都是人类自身特有的部分所不能做的，事实上的确如此。但如果真是这样，就存在着这样一种反过来的让人毛骨悚然的想法：我们是不是可以将一个人粗略算作是人类身体特有的一个部分，他是否可以等同于自己那个完整的、可正常运作的大脑？是你有脑子，还是你就等于你的脑子？这个问题并不好回答。砍下了我的胳膊，我仍可以签合同，用脚趾头拿笔或者发出一个口头指令，但如果把我的脑子弄傻了，那我的手再怎么做也不能算是在签合同了。摘掉了眼球我就再也看不见了，除非我可以安一个假眼球，这并不是一个遥远的幻想。但如果一块一块“取掉”你的大脑会发生什么呢？如果完整地保留下我的眼睛和视神经，而移除掉我的枕叶皮层，那我可能就会变成“皮质盲”，只残留下一些视力，著名的盲视就属于这种情况。毫无疑问，我们还可以多切除掉一点大脑，这样连盲视都不存在了，但你仍然活着。很多人都认为，这样在大脑上一刀一刀地割除下去，消除掉听觉、触觉、味觉还有嗅觉等，最后剩下的便是你的终极指挥部，那就是

你所在的地方，是决定人之所以是人的那个东西。

然而，这种观点是错误的。大脑的无数多种能力是纠结在一起、相互影响的，根本不存在一个将大脑所有的能力都集结于一处从而产生意识的中心。[30]让你成为你之所是的那些能力、性情、偏好和怪癖，都依赖于连通你的身体和大脑的通路才得以形成。哲学上有一个非常流行的有关脑移植的思想实验：你想成为哪一方？大脑的“捐献者”还是“收容地”？这只能通过非常扭曲的理想化手段来实现。如我以前所说，“你无法把我从我的身子上撕下来，只干净利索地留下一个轮廓”（Dennett, 1996a, p. 77）。

互联网有个最重要的特性：去中心化，走遍世界你也找不出它的中枢或者指挥部，所以你也无法对准哪个地方投一颗炸弹，一举把它销毁。计算机的每个部分都功能多样且沉积着海量的冗余信息，所以即使哪个部件出现问题，计算机也只会在性能上出现些许“适度的降级”。《2001：太空漫游》（*2001: A Space Odyssey*）中的智能计算机哈尔（Hal）却有它的“逻辑记忆中心”，里面满满地装着记忆卡，因此一旦戴夫（Dave）将这些记忆卡一个一个分开，哈尔就再也无法重新启动了。互联网没有类似这样的中心。尽管大自然没有赋予我们如互联网那般紧罗密布的分散结构，但不能否认，构成我们的各个部分也并非功能单一，它们的组织方式也是去中心化的。科学家们证明，我们的大脑组织是非常具有可塑性的，它能变换成各种新的配置，即使割舍掉了身体上的一些重要部分，只要不是那些致命的身体部分，你就还是可以做自己，追求梦想、惑乱敌人、权衡策略、重新尝试取得新的成功。这就是我为什么会说我们的确拥有自己身体各部分没有的能力。或者可以反过来说，要想理解或者弄明白一个完整活体中各个部分的能力，唯一的出路是你得认真思考它们是怎样为那个杰出整体的系统协调做出贡献的。

再来多看几个例子：你的大脑不懂英语，你懂；你的大脑找不出那个笑话的笑点，你能。纵然在理解和鉴别的过程中，是你脑中的相关结构在起着主要的因果推论作用，但离开了感官、四肢和其他器官长年累月的训练和支持，它也很难完成自己的工作。[31]

所以，将人这个生动的活体，这个有耐性、有意识、有理性的行动者记为主语并不只是我们的一种日常习惯：是你，犯了错误、赢得了比赛、迷恋上了莱斯利、说着凑合的法语、想去巴西、青睐金发碧眼的女人、犯了诽谤罪（在第62章中也有相关论述）。而我们之所以会饿、会累、会暴躁，所有这一切也只能归功于我们所谓的次人部分。

那些构成人类的部分到底是什么呢？它们是像砖头那样堆砌出一个活的人体吗？如果着眼点仅是原子这种最小的部件，那么事实的确如此。但如果在分子、细胞以及更高级的层面上，构建方式就不只是堆砌那么简单了。蛋白质是构成我们人体细胞的主力，它们的能力超出你的想象，就像无数微小的机器人，我们可以称之为纳米机器人。神经细胞更是万能，它们在大脑中主要负责传递、转换和调节工作，我们可以叫它微型机器人。在更大的结构中，神经元可以连接成不同的团体，这些团体之间相互竞争也相互合作。神经元来回地传递信息，彼此互相压制，它们可以分析从感官传来的信息流，唤醒“记忆中”休眠的信息结构（这些信息结构并不是大脑中单独的一部分），在你行动时，它们还会精心策划一些微妙的级联信号，让你的肌肉活动起来。

任何比原子直接堆积而成的结构更高级别的结构都展现出了行动者的特质。换句话说，它们可以被解释为意向系统。就分子来说，如动力蛋白、DNA校正读码酶、细胞膜上数以万亿计的通道蛋白等，它们干起活来几乎是“自动化的”，但依然能给人留下深刻印象。

举两个虚构的例子，就像《魔法师的学徒》（*The Sorcerer's Apprentice*）中结队行进的扫帚，或者像麦克斯韦妖（Maxwell's demon）。(32)再到细胞这一级，单个的神经细胞具有更强的试探性，它们不停地到处游走，寻找更好的神经链接，可以根据经历随时改变自己的放电模式。它们不再像是蛋白质纳米机器人那样的机器，倒更像是囚犯或者奴隶。你可以将它们想象成是正在狱中服刑的神经细胞，漫无目的地忙碌于大量连它们自己都搞不清楚的工程项目之中，对于变革政策以改变命运，它们从来都漠不关心。比细胞更高级别的结构单位，如神经束、神经柱、神经节、“神经核”这些细胞群，开始逐渐具有了一些深谋远虑，它们具有了自己专门的功能，可以敏感地感受到更多条件，甚至是来自外部世界的条件。它们看起来更像是行动者了，“工作起来”需要更敏锐的洞察力，甚至还需要做决策。

这些行动者像极了那些白领、分析师和高管，他们都有自己专门的职责。就像白领们一样，这些行动者同样有着健康向上的竞争热情，会力所能及地完成自己应该做的事情，在对手或者同伴失利时，它们甚至还会毫不犹豫地代劳。这个层次上的这些次人部分完全就是一些智能砖块，我们开始逐渐看出一个大体的轮廓，在这些次人结构的基础之上，究竟要怎样才能勾画出对一个人的整体理解。面对满满一箱自行车零件，我们不得不感慨还得要“组装一下”，但至少我们不必去裁截钢铁，

自己制造螺母和螺栓。

这些无脑的微粒是怎样构成一个活生生的人的？对于这个令人望而生畏的问题，我们可以掰开揉碎了看。可以自下而上，从细微处入手，就像我刚刚做的那样；或者也可以采取自上而下的方式，从整体的人开始，看看有哪些机灵的小东西一点点聚集起来，齐心合力地完成了能使一个人运动起来的全部工作。柏拉图倡导的是第二种方法，他将灵魂分成了具有类行动者性质的三种类型，相当于哲学王、卫国者和劳动者，或者是理性、激情和欲望。这并不是一个好的开始，对此两千多年来人们已经讨论得太多了。20世纪，弗洛伊德提出了本我、自我和超我三个概念，这无疑是一个进步。但是，真正打碎整体思维，指引人们进入次思维研究的是计算机的发明和人工智能研究的诞生。人工智能研究从一开始就确立了自己的目标——要把人类整体的认知能力转变成一个由次人类专业处理装置组成的巨大网络，包括目标发生装置、记忆搜索装置、计划评估装置、感知分析器和语句分析程序等。

20

大脑中的小人儿委员会
认知科学的好理念

在几千年以来的思考和探索中，理论家们总是会不经意间就将心智想象成一个存在于我们身体内部的行动者，就像一个小人儿，拉丁文称为“homunculus”，他可以坐在大脑控制室里完成所有精妙的工作。如果你将人类的神经系统想象成一个巨大的电话接线网络，就像思想家们在20世纪五六十年代喜欢去做的那样，那么对于网络中心的操作员，你可能就会有这样的疑问：他的心智是不是也是由一个更小的电话网络构成？这个网络是不是也有一个中心操作员？他的思想是不是也由……？任何一种理论，一旦做出了类似“中心小人儿”这样的设定，那么它就注定要进入一种无限的倒退。

但是，也许假定有一个小人儿并没有错，错误的是我们假定了一个中心操作的小人儿。在我的第一本书《内容与意识》（*Content and Consciousness*, 1969）中，我犯了一个严重的错误：没能抵御住说句玩笑话的诱惑。我这样写道：

> “大脑中的小人儿”被赖尔（Ryle）称为“机器中的幽灵”，它如今早已声名狼藉，因为它并没有对有关思维的种种问题做出什么贡献。尽管“大脑书写”这种类比会带来一些有益的启发，但它做的也只不过是将脑中的那个小人儿换成了一个委员会组织。（Dennett, 1969, p. 87）

那么，替换成一个委员会又有什么问题呢？（哈哈，这里用到了我的归谬法！）在书写《头脑风暴》（*Brainstorms*, 1978a）的过程中我逐渐认识到，用一个委员会去替代大脑中那个小人儿的想法其实并没有那么糟糕，我想，它应该是认知科学中一个不错的基础理念。下面是有效的老式人工智能（good old-fashioned artificial intelligence，即GOFAI; Haugeland，1985）的经典战略，人们称之为小人儿式机械主义：

> 因为人工智能程序员们是从一个带有意向色彩的起点入手的，所以他们当然会将计算机拟人化。比如说，一旦计算机顺利解决了问题，程序员就会说他们设计的计算机能够理解那些用英语写出的问题。程序设计的首要的、也是最高级的要求是要把计算机拆分成数个子系统，每个子系统都要执行一些带有意向色彩的任务。程序员列出一张流程表，上面贯穿着计算机评估、记忆器、鉴别器、监督器等部分，其中的每一个都算得上是小人儿……每个小人儿进而又可以被分解为更小的小人儿，值得一提的是，它们的智能性也会渐次降低。分解到最后，小人儿们都变成了一些加法器或者减法器，它们的智力仅能让它们从两个数字中挑出较大的那个。分解到这个程度，这些小人儿就可以由一台台机器替代了。（Dennett, 1978a, p. 80）

这一策略的独特价值在于它釜底抽薪地反驳了无限倒退论。小人儿式机械主义绕过了险恶的无限倒退论，用有限的倒退取代了它。就像前面已经说明了的，在倒退终止的地方，运行操作变得相当乏味，这时候就可以用机器来替代了。关键在于，以往在我们的印象中由中央操作执行的那些事物，现如今都已经被分担给了下属稍显次要、略微愚钝的那些行动者，这些行动者的任务还会继续被分配至更低层次的行动者，以此类推。

这是一个了不起的进步，但是有效的老式人工智能那种自上而下的工作方式要求有严格的等级分工，这却是个不必要的设计！我们可能已经想到要去除集权国王或者首席执行官，但还是有一大批副主席相互交涉，构成了系统的最高阶层，大量的中层主管要向他们汇报工作，同时这些中层也要指挥部下的行动，他们的部下也继而会召集更多的下属干事，如此等等。早期人工智能在计算机上的操作形成的就是这样一种超级高效的组织结构，它没有多余的动作，没有职位的闲置，不接受任何反抗。这在很大程度上是基于这样一个事实：运行那些早期的人工智能模型的大型计算机以今天的标准来说又小又慢，而人们总希望能更快速地得出结果。要想让你的人工智能系统赢得基金资助，就不能让它耗费很长时间才回答出一个简单的问题。它运行起来要非常高效。此外，书写成千上万行的代码是一项繁重的任务，因此如果你能成功地将目标任务，比如回答一些有关月岩的问题、诊断肾脏疾病或者下棋，分解成一组比较容易掌控的子任务，那你就可以清楚地看出：应当怎样编程、怎

样将程序整合成一个可行的系统。你可以有一套独到的“概念验证”(33)方法，它可能省时又省力。

值得注意的是，在设计计算机时，需求和工作表现几乎是完全独立的。硬件上，电力均衡地、源源不断地发送，没有哪个电路会面临电力耗尽的危机。软件方面，调度器系统总是让优先级最高的程序运行。至于哪个程序具有优先权，则需要通过某种竞争机制挑选出来，但各种程序是按照一个队列有序运行的，并不会陷入你争我夺的厮杀。正如马克思所说：“从各展其能到各得所需。”计算机科学家埃里克·鲍姆（Eric Baum）将这种等级控制恰当地称为“政治”。也许普通大众之所以会模糊认同这一事实，是源于一种普遍的直觉：计算机不会关心任何事。这与制造原料无关，我们凭什么说由碳元素构成的东西就要比由硅元素构成的更会关心什么事呢？而是因为，风险和机会的观念并没有被置入到计算机内部，所以它不需要关心。

而神经细胞则大不相同。构成人体的那些普通细胞与成群的工蚁比较相像，它们没有私心，机械地工作，只是些百依百顺的奴隶，干着一成不变的活儿，在一个不怎么有竞争性的环境中生存。但据我所知，大脑细胞必须参与激烈的市场竞争。为什么呢？神经元想要的是什么呢？就像它们真核单细胞的祖先，还有更远亲些的细菌和古生菌那样，神经元想要的是填饱肚子、保持能量，继续活下去。神经细胞类似于生物机器人，从任何意义上讲它都不能算是有意识的。请记住，神经细胞都是真核细胞，类似于酵母细胞或者真菌。所以，如果说单个的神经细胞有意识，那岂不是说脚癣也能感知！但神经细胞有些地方还是会像它们那些无脑的单细胞表亲：在生死攸关的争斗中，它们都不愧是有极强战斗力的行动者。它们的竞争环境不是在你的脚趾缝里，而是在大脑所要求的环境里。在这里，谁能够更加有效率地建立起联系网络，谁能够在足以识别出人类欲望和冲动的虚拟机层面发挥更大的作用，谁才有机会赢得胜利。

神经系统中的很多子系统都以敌对的方式存在，每个子系统都希望以自己的方式行事，彼此之间常常会陷入拉锯战。例如，我们的各种情感会为了取代对方而不遗余力地相互展开疾风骤雨般的对抗，有时候它们也会结成联盟来阻挠自己的劲敌开展行动。现在我渐渐明白，人类各种情感之间动态的对抗过程，以及情感对我们思想的影响，都是以协调单个神经元相互竞争天赋的神经化学系统为基础的。值得注意的是，不像是那些本质上就极为自私的行动者，比如癌细胞，从一个更广阔的运

作范畴来看，神经细胞还算是不错的团队成员。

我记得诺贝尔奖得主、生物学家弗朗索瓦·雅各布（François Jacob）曾说过，每个细胞都梦想着能完成分裂变成两个细胞，只有神经元不图繁殖，一心只想着要保持活力、有所作为。按照这种观点，对动物行为的智能控制是一种计算过程，就像在股票市场里完成一次交易那样。但正如神经科学家塞巴斯蒂安·承（Sebastian Seung）所说，神经元也是自私的，它们总是会倾其所能，最大程度地吸纳能在我们的大脑中获得的所有酬劳，这些酬劳以不同的“货币形式”出现。那么神经元利用多巴胺、5-羟色胺还有催产素都买了些什么呢？它们购买的是更加强大的影响力，能让自己在所处的连接网络中变得更加可靠。骡子虽不能生育，但它们仍会努力地营生。同样，神经元也有自我保护的本能，本质上，这是从它们可繁殖的祖先那里遗传下来的。

受神经科学的启发，作为大脑运作的一种模式，那种自下而上的小人儿式机械主义真是让人越看越觉得合理，因为从生物学的视角看，它生成的更加错综复杂也更有竞争性的“计算体系”是更合理的：我们可以看到它整个的发展历程，从胚胎直到成人，在这个过程中不断建构和修改。我们也可以看到，这种复杂的体系结构是如何从简单的神经系统发展而来，由那些未成熟的小人儿组合而成，那些小人儿有着“近似”于感知、标记和记忆的能力。[34]

21

近似算子
达尔文渐进主义的启示

为什么我要沉湎于这种对于“近似”（Sorta）的讨论呢？因为一旦要对各层级的功能运作做出分析或综合评价，我们就需要针对那个层级，追问这两个问题：它是什么？它都做些什么？“它是什么”是指其组成部分的结构组织，前提是这些组成部分都能如我们假设的那样发挥它们的功能。“它能做些什么”是指，它可以将一些近似认知的功能实现得很好，让我们能在一个更高的水平上安心工作，因为我们相信脚下踩着的这块基石一定可以很好地发挥自己的功能、被很好地利用。它为“思想怎样由物质机制构成”这一出奇复杂的问题中最难攻克的那部分点出了关键性的一笔。

其实早在计算机发展的黎明时期，当之无愧的计算机创造者阿兰·图灵就展望到了这一幕。他从不掺杂任何智力因素的机制片段入手，将它们组织成能力更强大的机制结构，这些强大的机制结构又可以进化得更为强大，如此下去，看不到尽头。在认知科学中，我们会将近似算子看作是达尔文进化论中渐进主义的一种平行理论（在第五部分中会有更加详尽的阐述）。有细菌之前会先有近似细菌，有哺乳动物之前会先有近似哺乳动物，有狗之前也会先有近似狗的动物出现，如此等等。

要弥补一只猿猴和一个苹果之间的巨大差异，我们需要借助达尔文的渐进主义理论；而要弥补一部人形机器人和一个手动计算器之间的巨大不同，我们需要借助的是图灵的渐进主义理论。构成猿猴和苹果的基础成分都大同小异，不同的是这些成分的构造方式，还有各个层级的功能对它们的开发利用。近似猿和猿在原则上并没有什么明确的划分。人形机器人和手动计算器也是由相似的基础部件组成，它们都是些无思想、无感觉的图灵程序块。把这些程序块视为组成部分，放到更大的功能结构中，再组合成更高层次的结构，以此类推，到最后我们能得到这样一些部件，它们拥有近似智能，能组合成被称为“理解”的那种能力。

而现在，我们从意向立场出发，开始追溯每一层级上近似理性行动者的信念和欲望，或者近似信念和近似欲望，从最简单的细菌到动物，从海星到天文学家，追溯组成大脑的识别、标记、比较和记忆回路。但即使追溯到人类这一层级，我们也没能明确地找到真正形成“理解”的那个点。当那位小朋友说“爸爸是一名医生”时，她近似明白自己所说的那句话，我也近似理解“$E=mc^2$”。某些哲学家抵制这类反本质主义的理论（见第43章）：你或者认为雪是白的，或者不那么认为；你或者意识到了，或者意识不到；没有什么所谓近似的心理现象；要么就是，要么就不是。对于这些思想家来说，思想的力量永远神秘莫测，它是“完美的”，绝不是什么仅凭物质机制就能造就出来的东西。

22

神奇组织
不要为逃避难题而求助于“魔法”

李·西格尔（Lee Siegel）有一部关于印度街头魔法的杰出著作叫作《魔法网：在印度的奇观和欺骗》（*Net of Magic: Wonders and Deceptions in India*），他在其中这样写道：

> 当我说“我正在写一本关于魔法的书”时，便会有人问：“是有关真正的魔法吗？”人们常常将真正的魔法理解为神迹、奇迹般的行为或者超自然的能力。我说：“不，只是些小伎俩，不是真正的魔法。”换句话说，真正的魔法不是真实存在的魔法，那些可现实操作的、真实的魔法不是真正的魔法。（Siegel, 1991, p. 425）

按照定义我们会说，“真正的魔法”不可思议，与自然法则相悖。很多人愿意相信真正的魔法确实存在。但魔术师、怀疑论者、驱魔师神奇兰迪（the Amazing Randi）却成功地复制了尤里·盖勒（Uri Geller）这些自诩巫师者的魔术，他想向世人证明，那些惊人的效果并不源自真正的魔法，它们都只是些戏法而已。不过，这并没有说服所有人。多年前，兰迪在温尼伯曾有过一次演出，在节目之后的问答环节上，一些观众便开始控诉他，说他其实是在双重欺骗：你和盖勒一样都是专业巫师，你拆穿盖勒的魔法，只是想让人们觉得，只有你才是真正的魔法师，这样就可以乘着盖勒的东风收获更多的名誉和财富。这些质疑莫名其妙，但兰迪却很难反驳，除非他能给在座的观众澄清这些戏法的原理。可兰迪尊重魔术界的世界惯例，他不愿意那么做。魔术师组合佩恩和特勒（Penn & Teller）曾经倡导要揭露魔术中的秘密机关，于是顶着魔术师同僚们的谴责，他们在魔术表演的同时，揭露出魔术背后的秘密。

就像很多人热切希望有真正的魔法存在一样，这种希望也折磨着很

多人对思想与大脑之间关系的考察。不少神经科学家、心理学家，甚至是哲学家，潜意识里还是会被这样的观念吸引：神经组织中的动态属性可以利用某种科学无法解释的潜在力量，做出一些让你觉得神奇的事情。也许他们的想法没错。但我们不能一开始就这样假设，我们的前提条件必须是：没有什么神奇的组织！

有些东西是可以绝对肯定的：在计算机程序中，没有什么东西是在物理上无法解释的；没有什么难以想象的力场，也没有什么神秘的量子论诡计，更没有什么“生命冲动”（élan vital）。不可否认，计算机中没有神奇组织。我们清楚地知道那些基础任务在计算机中是怎样完成的，我们也知道它们是怎样一步步地组合成为越来越复杂的任务的，在解释这些功能建构时不需要添加任何神秘色彩。现如今，对于计算机表现出的那些精湛技术，我们常常是赞不绝口，但尽管如此，作为一部机器，计算机对我们来说和一个开瓶器并没有什么差别，都一样平凡又熟悉。它里面没有什么“真正的魔法”，都只是些戏法！

这一事实非常重要，我们在下一章还会详细讨论。它的价值在于：无论计算机做出了什么看似神奇的事情，我们都有证据表明它无需借助神奇组织。也许大脑采用的是另外一种运行方式，也许它真会用到一些神奇组织，也许兰迪和盖勒都是真的魔法师，但我们还没有找到合适的证据。所以说，计算机在清楚解释事物方面扮演着非常重要的角色，不管是飓风、房地产泡沫、艾滋病病毒，还是人类意识，在理解事物时，我们有充足的理由要首先去开发它们的计算机模型。

神奇组织这一概念作为一件思考工具，它的功能类似于警察的警棍：你拿警棍惩罚别人，警示他们不要使用不正当的理论。当然，就像你可能滥用警棍一样，你也可能滥用这个概念。神奇组织这件思考工具与“奥卡姆剃刀”有相似的作用，所以它同样会引发一种系统式的保守主义，令人缺乏开创性和前瞻力。现代遗传学的创始人之一威廉·贝特森（William Bateson）清楚地说明了这一点，我很喜欢他的这个例子，以下就是他在1916年说过的那段话，时间还没有过去很久：

> 生物的特性在某种程度上是依附于它的物质实体的，或者也一定程度地取决于核染色质（染色体）。神奇的是，不管这些物质实体或者染色体粒子有多复杂，它们都拥有传递我们的特征和种族（基因）的力量。我们可以推测，染色体的粒子虽然彼此无法区分，几乎在所有已知的测试中都是同质的，却可

以凭借自身的物理特性赋予新生命所有的特性，其能力远远超出了最深得人心的唯物主义理论的解释。

当然，贝特森想不到还有DNA的存在，人体的每一个细胞中都含有包含着三十亿个碱基对的双螺旋，这超出了他的想象。而幸运的是，不是所有的生物学家都如贝特森一般悲观，他们始终在研究，遗传信息是怎样通过一些神奇的微粒代代相传的。在探索道路上，这些生物学家们一直坚持“没有神奇组织”的原则。从遗传学角度看，对于自己研究对象所具备的各种能力他们早已了如指掌，剩下的任务是尽可能地做出一个物理模型，让它也同样具有类似的能力。

如今，我们也面临着同样的任务。实验心理学已经为我们提供了一张有关心灵能力和弱点的十分详细的列表，其中显现出了成功的知觉和失败的错觉，说明了语言学习的步骤，还包括造成注意力分散、欲望、恐惧及欢乐的各种条件。而现在，作为“深信不疑的唯物主义者”，我们需要向大家解释清楚大脑是如何做到这些的，不能借助神奇组织。

随着理解的不断深入，我们对神奇组织的看法也在改变。20世纪80年代中期，“联结主义”和其他一些“神经网络”模型的突然出现[35]展现了小束神经元的学习能力和图形识别力，而就在几年前，这根本就是无法想象的。虽然我们仍然不知道大脑是如何产生这些理论模型所展示出的计算能力的，甚至不知道大脑是否真的是像这些理论模型描述的这样工作的，但至少我们现在可以假定神经网络有某种连接能力，只要设定的这种能力不超出应有的界限范围。作为前提，我们无需对此多做解释。兰迪不可能完全照搬盖勒的魔术，但我们可以就此得出结论，兰迪的表演多多少少可以解释盖勒的伎俩，这就给我们更深入地探索实际发生的过程指明了方向。反对神奇组织的主要理由是：它提供给我们的是一种逃避问题的方法，而不是一种解决问题的方法，因为神奇组织预设了这是一个根本无法解决的神秘问题。

23

身陷机器人控制室
初长成的大脑不是一块“白板”

根据定义来说，机器人没有神奇组织，所以它们可以为思想实验提供一个纯净无干扰的平台，我们可以举出这样的例子。

一天清晨，你醒来后突然发现自己身处在一个没有窗户、非常诡异的房间里，躺在一张从没见过的床上。屋内有两面装满了不同颜色的微型闪光信号灯的墙，另外两面墙上则布满了成千上万的按钮。信号灯和按钮上都写着数字，但没有标记出哪个与哪个相对应。在床边的桌子上，有人给你留下了一张字条：

> 早上好啊！我们麻醉了你，在你睡着的时候把你绑架到了这里。冰箱里有食物，角落处是盥洗的地方，在这里，你身体的各种需要都能得到满足。你现在是被囚禁在一个巨型机器人的控制室里，每盏亮起的灯都说明了大量有关机器人状态的信息，这些信息是由高度复杂的神经网络分析器根据机器人的高清视频眼、麦克风耳、触摸传感器和嗅觉传感器摄入的原始资料分析得出的。按钮控制着机器人的行动，已经调试好，随时可以使用。
>
> 机器人所处的环境非常险恶，危机重重，但也存在机遇。你掌握着它的命运，与此同时，如何控制这个机器人在世界中生存也决定着你的未来。机器人一旦毁灭，你房间里的电力系统也会全线瓦解，电冰箱里不再会有食物，你将面临死亡！祝你好运！

这真让人恶心！你提心吊胆地开始按动按钮，每按一下都要看看有什么事情发生。你按下了编号为4328的黄色按钮，随即，那盏编号为496的蓝色信号灯渐渐熄灭了。这是在让机器人挠痒痒吗？或者是让它

闭上眼睛？又或者是让机器人“吃”了些东西，补充它新陈代谢的需要？你第二次按下4328按钮，分散排布的几盏灯亮了。世界会因此发生什么呢？这到底意味着什么？你焦躁不安，你知道这些信号灯蕴含的信息数量庞大，但却不知道哪盏灯与哪条信息对应，也不知道每个按钮都控制着机器人的哪些动作。

要是信号灯上和按钮上都有标注该多好！要是标注用的语言你还都能看懂，那么你就可能摆脱这些困扰。又或者，如果这间屋子有扇窗户，打开它能看到外面，那么在按下按钮后，你至少能看到外部世界相应的变化。这个机器人有胳膊和腿吗？假如有了这扇窗户，你还能试着将外面世界的活动与墙上闪烁的信号灯相互联系起来。而没有这扇窗户，尽管所有的信息也都摆在那里，但它们却不能构成任何解释。你可以让机器人做成百上千个动作，但却无法搞清楚这些动作对世界造成了什么影响。

现在，你进退维谷。因为，不管你有多聪明、多有想象力，即便所有信息资料都任你取用，你也很难说出发生在墙面上的这些到底表示着什么。但是如果你无法解决这个问题，那就很矛盾了，因为现在的困境不也正是大脑所面临的困境吗？大脑也是被困在了一个没有窗户的房间里，也就是你的头盖骨，成千上万的输入线路在忙碌着为它传入外部世界和你身体状态的相关信息，成千上万的输出线路在淡定自若地操控着肌肉的收缩和舒张。大脑当然不会在头盖骨上凿出个窗户，朝外看看周围发生了什么事情。凿开这么一扇窗户到底能让你的大脑得到什么好处呢？它跟你不一样，它没有自己的双眼，只能收容你眼睛接受到的信号或处理过的信息；至于这世上各种事物的样子，它也没有什么先天的记忆。

有些人并不觉得大脑的工作原理有多难，因为在看一只鸭子的时候，我们视觉皮层表面的激活模式，也就等同于那些闪着光的信号灯，就会形成一个鸭子形的图案。(36)好吧，但如果你的大脑和你一样，通过学习认识到了鸭子的模样，那请问，它到底是怎样学习到一只鸭子的模样的呢？

如果信号需要“解码”，你的大脑该怎样才能认识到事物呢？解码后的那些信号都变成了些什么？土耳其语？如果你不懂这门语言的话，土耳其语标签对你来说就毫无意义。所以，要从输入资料那里得到些有价值的信息，我们的大脑是不是得先懂得一门语言？之前有过这样一种观点，大脑无需后天学习，它自身就带有一种内部语言：心理语言

（Mentalese），也就是思想语言（Fodor，1975，2008）。看起来，这个撩人的观点似乎能带领我们往正确的道路上再迈进一步，但是，除非能了解到它最初的成形过程及运作的诸多细节，否则“存在一种内部语言”这个观点只是给问题换了一个名字，并没有将问题解决掉。

我们知道，大脑总是在以某种方式解决问题，它可以果断地为输入信息造成的那些窘况或者机会找出适当的对策并且输出。我们也知道，大脑解决问题的方式绝对不是使用了一门语言，比如英语或者土耳其语，因为这种解决方式并不是像我们小时候学习母语那样学到的。或许大脑内部语言更像是书面语，由记忆在大脑存档中进行书写？又或者，它更像是纯粹的口语？它会有成千上万甚至上亿个“词语”吗？词语的排列顺序和意义有关吗？它有语法吗？大脑的某一区域能够准确地理解其他区域中的信息吗？

如果思想语言指的是封闭控制室里的一个小人儿所明白的那种语言，就像在机器人体内的操纵室里，你能读到信号灯和按钮上的标签一样，那么实际上，它是把“如何在由毫无理解力的部分组成的机制的基础上完成理解和学习过程”这一问题延宕了，根本没有给出回答。而如果将小人儿从思想语言的假设中取消掉，那么不管整个这套系统是什么，都跟语言没有多大的相似度了。而且到目前为止，还没有人对“思想语言是如何运作、如何在发育过程中产生的”这些问题给出过详细的说明，所以，也许我们不该再那样自欺欺人，不该再那样自我安慰地说，我们在这个方向上正在取得进展。其实我们一直都在原地呢！

这个直觉泵是想说明，困在控制室里的大脑和那时受困的你所面临的境况并不全然相同。对于大脑来说，输入的信息是与输出相连接的，大脑本身就拥有一些能力，知道要去认识什么，知道怎样才能使各种连接变得更加合理。这项任务是提前完成了的。这也以另外一种形式证实了那个广为人知的说法，初长成的大脑不是“白板”一块（Pinker，2002），它已然经历了自然选择的雕琢，自身体现着各式的偏好、预设和衔联，而且，有些合适的连接是提前内置的，无需标注。

在理解产生之前，先要有一些不需要理解的能力出现，这是自然法则。细菌浑身都是超凡的能力，这些能力让它们过得很好，只是细菌无须去理解这些能力，它们是愚蠢的。树也有趋利避害的能力，但它们亦不必知道其中的缘由。自然选择的车轮滚滚向前，一路上造就出了无数精巧能干的事物，这个过程也不包含任何理解。

我们成年人类具有的那种理解能力是进化史近期才出现的一种景

象，它是由一些结构组合生成，而这些结构所具有的能力应该也是由鱼或者蠕虫所有的那种类似理解力发展而来的。大部分时间里，这些结构发挥着应有的作用，但它们并不需要知道自己为什么会那样发挥作用。

而认为有一名发育健全的理解者被关在控制室里，面对着墙面上各种各样的输入和输出信息，这种观点注定是行不通的。为什么？因为如果这种理解力无法解释，那么我们就得建立一个神奇组织、一种神秘力量作为理论的基础。但是，如果这种理解力的运行过程是可以清楚解释的，是由一些不具备理解力的过程、活动和能力构成的，那么你又回到了解释理解力如何从其他能力中发展而来的道路上。既然如此，又何必绕这个圈子呢？

第三部分

关于计算机的思考工具

INTUITION PUMPS AND OTHER TOOLS FOR THINKING

也许你们已经发现，我前面已多次提到计算机，而在后面的章节，我还会写到更多有关它的内容。今天，计算机无疑已经成为了我们最有力的思考工具，这不只是因为它们承担起了单调繁重的苦差事，解放了智力劳动，而且，计算机科学家们发明出的许多概念也凭借着自身的实力成为了极好的思想工具。

如今，我们整日都在计算机术语（例如硬件、软件、带宽或者千

兆赫）的浪潮中浸泡着，所以我们想当然地就会认为自已对这类专业术语的理解绝对准确可靠。然而，直到在课堂上真正列出了几个这样的词语时，我才发现，虽然每个孩子也都装模作样地点着头，但他们对这些词的理解程度却实在是参差不齐，有时这些孩子们甚至会将我所传达的意思曲解到几近诡异的地步，以致我自己都被弄糊涂了。那么接下来，我首先想要教给大家的是，怎样为世界上最简单的计算机编写程序。

现在，如果你肯花费时间和精力去掌握一点基础的计算机编程技巧，那么你将会对我下面所说的这些有更深刻的体会，而如果你现在已经是经验丰富的计算机专业人士，那就希望我解释问题的方法能在你向普通人说明问题时发挥一点儿作用，或者跳过这一段也可以。我曾尝试对数百名有“计算机恐惧症”的本科生讲授下面的内容，效果不错：即使是那些宁肯背诵一整页电话号码也不愿意解决一道难题的人也不得不承认，能让一台超级简单的计算机按照我们的意愿做事是一件让人极有成就感的事。而当你完成了下文中的所有练习后，你就将掌握有关计算机能力的7个秘密。

24

计算机施展魔法的7个秘密
部分之和大于总体的最好实例

计算机隐藏的那种巨大能量一定会被几个世纪之前的人们看作神迹，看作“真正的魔法”。不过，尽管很多计算机程序都复杂得吓人，但它们全都是由一些基本的步骤组成的，仅凭几个简单的术语我们就能把这些步骤解释清楚。计算机上的运算从不夹杂任何神秘成分，作为思考工具，这正是它的重要价值之一，而且解释这件事儿本身就是一个很有趣的哲学问题。到底计算机是怎样施展它的“魔法”的？能在一个比较基础的层面上理解这个问题非常必要，这一章会提供给我们一些相关启示。

先从一台我们能想到的最简单的计算机——一台寄存器机开始吧，看看它有什么能力以及为什么它会有这样的能力。然后，我们再去考察图灵机和冯·诺依曼机。其实它们也只是一些更高效率的寄存器机而已，笔记本电脑能做的事情寄存器机也都能做，只是你得用它算上几个世纪。之后，我们才能慢慢地搞清楚，更高级的那些计算机“架构”如何在寄存器机这台本体机上增长了速度、壮大了技能，而人类大脑正是其中最重要也是最有趣的一种架构。

等一下，我从没说过“大脑就是一部庞大的计算机”，没有，一次都没有。我想说明的是，如果大脑真的是一部庞大的计算机，那么我们就一定能够找到一种方法将大脑所有的活动都解释清楚，不留下任何神秘色彩。我们采用的这种方法就是逆向工程：探究一个复杂的系统，分析它能做什么，是怎样做到的。逆向工程让我们知道了，心脏是如何像一个泵那样履行自己的职责，肺又是怎样收集氧气排放二氧化碳的。神经科学是逆向工程应用于大脑研究的一种尝试。我们都知道大脑能做什么：预测、支配、记忆、学习，接下来还需要我们搞清楚的是，大脑是怎样做到这些的。

这是个充满争议的话题。2000年，小说家汤姆·沃尔夫（Tom Wolfe）曾在自己的文章《抱歉，你的灵魂已死》（*Sorry, But Your Soul*

Just Died）中详细描述过这一敏感话题，随即便引起了一场激烈的论辩。如果我们不想把时间都浪费在那些雄辩和谴责上，而是要真正探索这个艰险的领域，我们还需要一些更锐利的思考工具。我们的大脑是否真的具有并利用了计算机永远都无法企及的那些神奇能力呢？要想有根据地解决这个问题，我们首先得知道计算机都能做些什么，它们是怎么做到的。而要想圆满地证明“我们的大脑不是，也不会是一台计算机”，可用的方法是：（1）证明大脑某些“活动部件”所参与的信息处理工作是计算机无法模仿的；或者（2）证明我们平日熟悉和喜爱的那些精神壮举并不能由组合、汇总计算机部件的简单工作或通过任何计算机形式的处理方式来实现。

除了哲学家外，还有神经科学家、心理学家、语言学家，甚至物理学家都认为这种人类大脑的“计算机隐喻”极具误导性，因为大脑显然能够做很多计算机不能做的事情。他们通常（但也不绝对）先对“计算机是什么”或者“应该是什么”这个问题规划出一个很天真的预设，最后就得出了这个明显但又不相干的命题，即大脑可以做很多笔记本电脑不能做的事情，因为笔记本电脑没有充足的传感器和效应器，内存低，运行速度也有限。而事实上，关于“通常情况下计算机能做什么”这个问题，如果真要回答，我想我们需要弄清楚的应该是：在通常情况下，计算机的能力来自哪里，是怎样得以发挥的？

1957年，那还是计算机时代的初期，王浩发明了寄存器机这个神奇的东西。他是逻辑学家，是库尔特·哥德尔的学生，顺便提一句，他也是一位哲学家。寄存器机是一件简洁的思考工具，每个人的工具箱里都该有这么一件，可惜它至今还仍未被大家所熟知。(37)寄存器机是理想化的、假想出的计算机，它只是由有限数量的寄存器和一个处理单元构成的，具有十足完美的合理性。

寄存器是一些存储单元，每个寄存器都有自己各自的地址，比如寄存器1、寄存器2、寄存器3，以此类推，它的内容包含一个整数，如0、1、2、3，……。寄存器就像是一个一个的大盒子，里面可以放置任意多的（从0到无限多个）豆子。不管这些盒子有多大，而我们总是希望这些盒子能无穷大，因为那样它们才能装上无限大的整数。但我们对大盒子的这些设定是有目的的。

处理单元只具备三种简单的能力：它只能“遵循”三种“指令”，一次执行一条。每一串指令构成的序列都是一个程序，而每一条指令由一个数字来标记，即给每条指令分配一个ID号。这三条指令分别是：

结束。表示将自己停止或者关闭。

增量。将寄存器n中的数字加1，或者说在盒子n中再填入一粒豆子，然后进入下一步，步骤m。

减量。将寄存器n中的数字减1，或者说从盒子n中拿走一粒豆子，然后进入下一步，步骤m。

减量指令与增量指令的运行原理基本相同，只是减量指令中多了一个非常重要的难点：如果寄存器n中的数字是0该怎么办？我们已经无法再从中减1，因为寄存器中不能持有负的整数，正如你无法从一个空盒子中再拿走一粒豆子，面临如此窘境只得另寻它路。办法是：跳转到其他地方运行下一步指令。每一个减量指令都需要在程序中列出那个接受跳转的位置，以应对当前寄存器数字为0的情况。所以，减量的完整定义应该是：

减量。将寄存器n中的数字减1，如果可以执行则继续进行步骤m，如果寄存器n不能减1，则跳转至步骤p。

一部寄存器机能做的所有工作，简单表示就是：结束、增量、减量（或者跳转）。

乍一看，你会觉得这台机器的工作实在有些无聊，只做着一些把一粒豆子放进盒子里或者从盒子里拿出一粒豆子的工作（只要它能找到一粒豆子的话，如果找不到，它会跳转到另一条指令）。但事实上，它可以完成一台计算机能做的所有运算。

让我们从简单的相加开始吧。假设你想让寄存器机将某一个寄存器（比如寄存器1）里的数字加到另一个寄存器（寄存器2）里的数字上。如果寄存器1中的数字是“3”，寄存器2中的数字是“4”，那么，我们要设计出的程序就是：相加结束后，寄存器2中的数字为“7”，因为3+4=7。这个程序用简单的RAP（Register Assembly Programming，寄存器汇编程序）语言编写出来就是：

程序1：相加［1,2］

步骤	指令	寄存器	下一步	跳转至步骤
1	减量	1	2	3
2	增量	2	1	
3	结束			

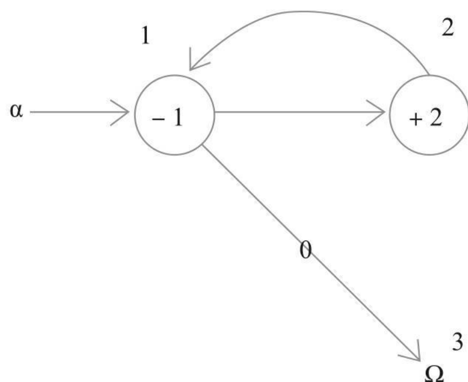
前两个指令构成了一个循环，寄存器1每减一次，寄存器2就加一次，直至寄存器1中的值变成0，处理单元能“注意”到寄存器中数值为0这一点，继而会将运行步骤跳转到3，即寄存器机停止工作。处理单元并不知道寄存器中的数字，它只能识别出寄存器中数字为0的情况。用盒子和豆子来打个比方：把处理单元想象成是一个盲人，它看不见盒子（寄存器）里面有多少豆子，但是通过探测，它可以识别出盒子空了。所以，尽管事实上处理单元不知道寄存器里确切的内容，但只要步骤1开始运行，寄存器1的内容（不管寄存器1中的数字是几）就会累加到寄存器2里（不管寄存器2中的数字是几），然后停止。（你能知道它为什么会这样运作吗？多试几个例子看看。）我们甚至可以这样说：即使不知道哪两个数相加，不知道数字是什么，甚至也不知道加法是什么，寄存器机也能完美地完成累加计算！

练习1

a．运用程序1，想一下寄存器机需要几步可以计算出2＋5＝7？（“结束”也算是一步。）

b．要计算出5＋2＝7需要几步完成？（你从中可以得出什么结论？）(38)

有一种图解可以很好地说明这一运算过程，我们称之为流程图。圆圈中的数字代表将要接受运算的寄存器的地址（它不是寄存器中的内容），“＋”代表增量，“－”代表减量。程序从α（alpha）开始，至Ω（omega）处停止。箭头表示进入下一指令。请注意，每个减量指令处都会向外引出两个箭头，可以减量时有箭头指向一个步骤，而无法减量时（当寄存器中的内容为0时，遇0跳转）则有另一个箭头指向另外一个步骤。

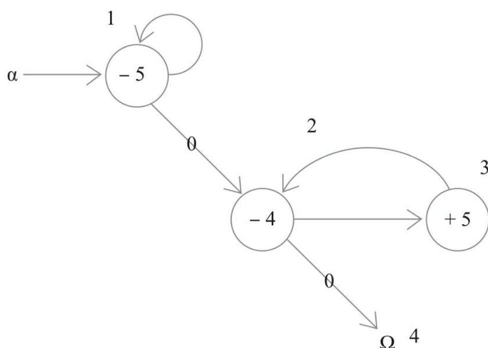


现在，我们就来编写一个程序，将一个寄存器中的内容移动至另外一个寄存器中。

程序2：移动［4,5］

步骤	指令	寄存器	下一步	跳转至步骤
1	减量	5	1	2
2	减量	4	3	4
3	增量	5	2	
4	结束			

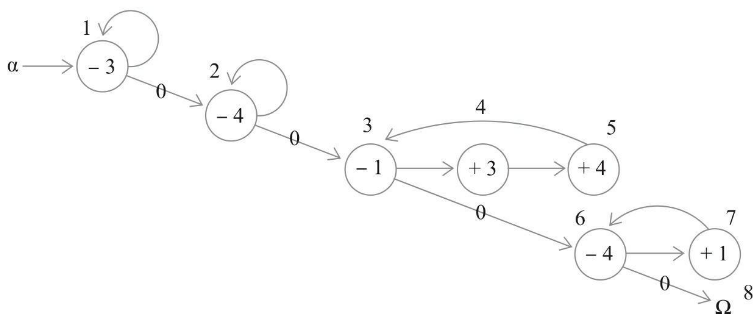
流程图如下所示：



可以看到，第一个循环是清空寄存器5。这样一来，不管一开始寄存器5中的数字是多少，它都不会影响到在第二个循环中要移进来的内容——第二个循环正是增量运算的环节，将寄存器4中的内容添加到已

清空的寄存器5中去。这种初始化步骤被称作寄存器清零。它是一项非常有用的标准化运算，我们会不断用到它，好让寄存器能重新投入使用。

第三个程序是，简单地将一个寄存器中的内容复制到另一个寄存器中，被复制的那个寄存器中的内容保持不变。仔细观察这个流程图，思考它的程序设计：



程序3：复制［1,3］

步骤	指令	寄存器	下一步	跳转至步骤
1	减量	3	1	2
2	减量	4	2	3
3	减量	1	4	6
4	增量	3	5	
5	增量	4	3	
6	减量	4	7	8
7	增量	1	6	
8	结束			

显然，这是完成拷贝的一条迂回路线。首先，我们要将寄存器1中的内容移动到寄存器3中，同时比照3中的内容做一个副本，将它复制到寄存器4中。最后，再将寄存器4中的内容重新移动到寄存器1中——也就是将寄存器4中的内容复制回了寄存器1。程序就这样按部就班地运作，不管一开始寄存器1、3、4中的内容是什么，到程序结束时，寄存器1中一定保留着自己原来的内容，复制的内容保存在了寄存器3中。

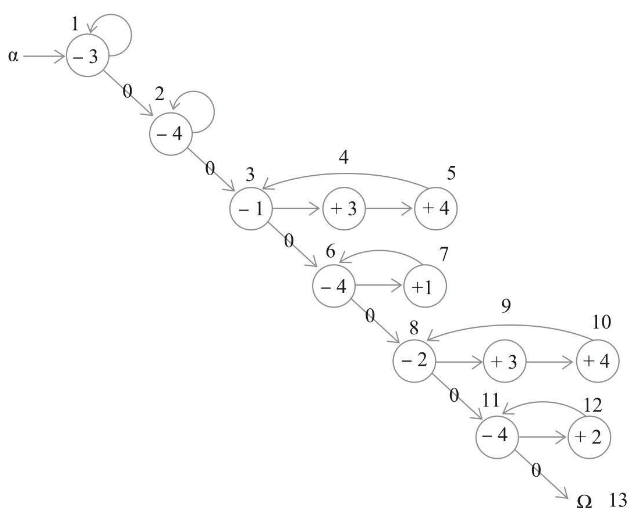
如果你认为自己对这个程序的运作方式还不是很清楚，那么你可以尝试“手动模拟”整个过程。取出几个杯子充当寄存器（用铅笔在每个杯

子上写一个数字，这是它们各自的地址），再找一些硬币或是豆子。在每个寄存器里都放置几枚硬币，然后记录下寄存器1和寄存器3中硬币的数目。接下来，依照程序按部就班地执行完毕，你会发现，寄存器1中硬币的数量与先前一致，而寄存器3中也拥有了与寄存器1数目相同的硬币。能将累加寄存器的基本运作过程都吸收内化是非常有必要的，我们得做到无需再去费劲思量它们。那么现在，就让我们先用几分钟去变成一部寄存器机吧（就像一个演员要去饰演哈姆雷特那样），因为接下来还要用到这种技能。

我发现好多学生常会犯这样的错误：他们认为，在做寄存器减量时，硬币从一个寄存器里取出之后就一定得放入到另一个寄存器中去。这是不对的。减量取出的硬币只需放回到那一大堆硬币中去即可，在简单的加减法套路中，有“无限多”的硬币可供你使用。

掌握了移动、复制还有归零这些技能之后，我准备重新回到之前的加法程序，对它进行改进。之所以要这样做是因为，程序1虽然成功地将加法运算的结果存进了寄存器2中，但在运行过程中，寄存器1和寄存器2中原有的内容却受到了破坏。我要设计出一个更加专业的加法程序，它能保留住原有的那些数值以备后用，将得出的结果放到另外一个地方。现在就让我们来思考这个程序，将寄存器1的内容与寄存器2的内容相加，两个寄存器中的数值保持不变，将结果显示在寄存器3中。

下面是完成这一程序的流程图：



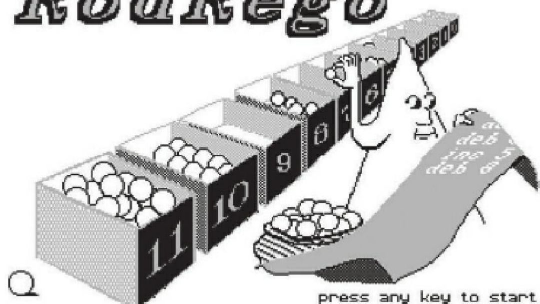
现在，我们就来分析一下各个循环，看看每一步是怎样运行的。首先我们要将显示最终答案的寄存器3归零，同时清零那个空闲的寄存器4，让它充当临时的接受器或者缓冲器。然后，我们把寄存器1中的内容复制到寄存器3和寄存器4中，随后将寄存器4，也就是缓存器中的内容再次移动到寄存器1，让它恢复原有的数据；在这个过程中，作为缓冲器的寄存器4再次被清零，以备后用。接下来，我们将上一步的寄存器1替换为寄存器2再执行一遍，将寄存器2的数据加到寄存器3的数据上。待整个操作完成后，缓存器（寄存器4）清零，加法答案显示在寄存器3中，同时，我们选用的两个加数也保留在了原来的出处——寄存器1和寄存器2里。

将这套由13个步骤构成的RAP程序改写成下面的流程图，可供处理单元读录：

程序4：无损加法［1,2,3］

步骤	指令	寄存器	下一步	跳转至步骤
1	减量	3	1	2
2	减量	4	2	3
3	减量	1	4	6
4	增量	3	5	
5	增量	4	3	
6	减量	4	7	8
7	增量	1	6	
8	减量	2	9	11
9	增量	3	10	
10	增量	4	8	
11	减量	4	12	13
12	增量	2	11	
13	结束			

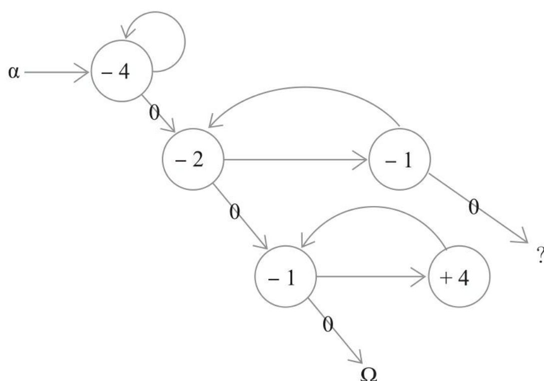
我不建议你拿杯子和硬币去手动模拟这个程序。人生苦短，一旦你把这些基本步骤都刻入脑海，那就可以切实享受一下“义肢”给你带来的便利了，这副义肢叫RodRego，你可以从网站http://sites.tufts.edu/rodrego/上下载到这部寄存器机。



最初的寄存器机RodRego的网站首页，1986年。

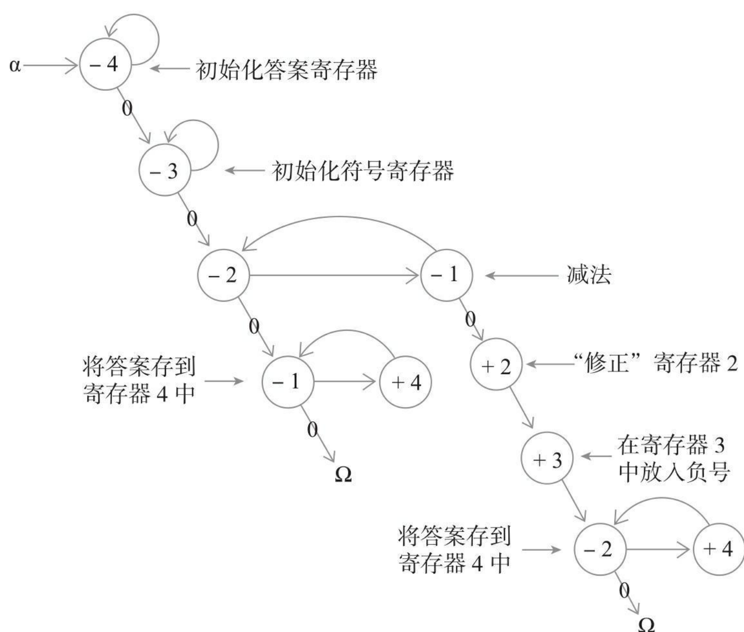
RodRego有PC、Mac(39)两个版本可在您的计算机上运行。20多年前，我和乔治在课程软件工作室里研发了这件思考工具，如今，成百上千的学生和其他行业的人员通过它成为了熟练的寄存器机思考者。你可以将自己编写的RAP程序键入电脑，观察它们如何运行，不管寄存器中的内容是数字还是豆子。另外，有的多媒体软件还可以顺着流程图将处理单元的程序路径淋漓尽致地展现出来，让你可以清楚地看到RAP指令与流程图各环节之间的相互对应。

接下来，我们再来看看减法。这是我的第一次尝试，从寄存器1中减掉寄存器2中的内容，将答案显示在寄存器4中。你能看出我这个流程图存在的问题吗？



这个流程图只有在寄存器1中的内容大于寄存器2的内容时才能正常工作。但如果实际情况不是这样，又该怎么办呢？寄存器1在不断减量的过程之中就已被“清零”，那接下来会发生什么？此时我们不能中止电脑运行，因为那样的话，寄存器4中会留下一个错误的答案“0”。事实上，在寄存器1归零时，我们可以开启一个新的程序，它会先倒退半个

循环，将寄存器2中最后减去的1再加回来。此时，将寄存器2中（而不是寄存器1）的数字转化成负数，所得的结果就是正确答案。所以我们只需把此时寄存器2中的数字移动至寄存器4中（之前，其中的内容已清零），再将答案标注为负数即可。我们还需要另外一个寄存器，比如说寄存器3。与寄存器4一样，寄存器3也要首先清零，然后，通过程序设计，它要像一面“旗帜”那样为答案做标记，内容为0时表示正号“+”，为1时表示负号“-”。下面就是这个程序的流程图，各个循环和步骤都有相关注释。（这些注释也可以写入RAP程序中供你以及其他人参考阅读，你要用“#”在其首尾做出标记，这样RodRego就会自动忽略掉它们）。



练习2

a．请为这个流程图写出它的RAP程序。（请注意，由于程序有不同的分支，你可以采用不同的顺序为各步骤编号，只要“下一步”对应的步骤编号填写正确即可。）

b．如果从3中减3，或者从4中减4，这个程序会出现哪些不同？

c．将寄存器3在第三步之前就清零而不是推延到完成第四步之后，这样会避免哪些可能的错误？

熟练掌握了加法和减法的编程以后，再设计乘除运算就易如反掌了。n乘以m表示m个n相加。我们可以这样来设计程序：用一个寄存器作为计数器，从m递减到0，每完成一次加n的运算就减去1。

练习3

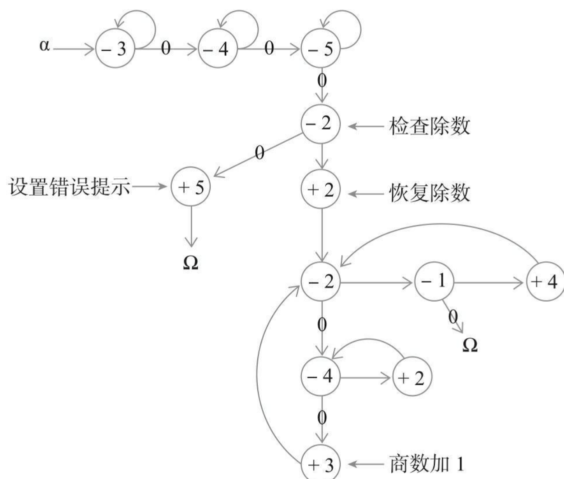
a．画一个流程图或者写一个RAP程序，让寄存器1中的数字与寄存器3中的数字相乘，将答案显示在寄存器5中。

b．（可选）[40]利用复制和移动改进你上面写出的乘法编程，使寄存器1与寄存器3在程序停止后都能保持原有的内容，方便你在运算完成后清楚地检查输出和输入的内容正确与否。

c．（可选）画一个流程图或者写一个RAP程序，要求它能识别寄存器1与寄存器3中的内容（同时保证这些内容不受损坏），将内容较大的寄存器地址（1或3）显示于寄存器2；当两个寄存器中的内容相等时，寄存器2中显示2。（程序完成后，寄存器1和寄存器3中的内容要保持不变，寄存器2显示出所含数字较大的寄存器地址，当内容大小相等时，寄存器2将报出数字2。）

与乘法类似，除法运算只需要从被除数中一次一次不断地减去除数并记录下相减的次数即可。如果有余数，我们可以留出一个专门的寄存器储存它。另外，我们还要小心地添加一个必要的检查项目：看看除数是不是0（0不能作除数，请问这是为什么？）。所以，在进行除法运算之前，我们得对除数作一个简单的核查——看看它能否被减量。如果可以，我们还必须再安排一次增量，以便使除数恢复原值，让除法运算继续进行。如果在进行减量核查时，除数真的是0，那就需要拉响警报了。为此，我们可以专门准备一个寄存器负责报错：寄存器5里如果有1出现就表示，“快！我就要除以0了！”

下面就是一个除法流程图，用寄存器1中的内容除以寄存器2中的内容，将答案写入寄存器3，余数显示在寄存器4，醒目的寄存器5用来提示“错误信息”（显示1表示“我将要除以0”）。



顺着这幅流程图梳理，你就能看到“除数为0”是如何中止掉整个运算的，报错的旗帜究竟是如何竖起的。另外还需要注意的是，寄存器4身兼双重职责，不仅要复制并恢复除数数据，以便下次继续成功地减量，而且，还要时刻准备着成为一个余数寄存器。如果在寄存器4将自己的内容倾倒回寄存器2以备下次减量之前，寄存器1中的内容就已经为0，那么留在寄存器4中的数字就是余数，它也正好显示在了余数寄存器中。

秘密1

发挥能力，无需理解力：有些东西——如一台寄存器机——它能完成精密的运算，但无需理解自己在做什么。

寄存器机没有头脑，因而它不能理解。但因为一碰上增量、减量、停这三条“指令”，它就会老老实实地执行，所以我们也有理由说，寄存器机是近似于明白这三件事情的。不过这三条指令不是真正的指令——它们只对于我们来说像是指令，但对于寄存器机，它们只是近似于指令而已。

也许你已经看到，减量跳转分支对寄存器机起着至关重要的作用。正是这个指令让计算机（近似）“注意”到了这个世界，继而它开始用这些注意到的事情指导自己之后的“行为”。其实这种条件跳转分支对所有存储程序计算机都是至关重要的。埃达·洛夫莱斯（Ada Lovelace）早在19世纪就认识到了这一点，这从她关于查尔斯·巴贝奇（Charles

Babbage）的分析机的精彩论证中就可以看出。而分析机正是所有现代计算机的原型。(41)

其实，只要掌握一定的诀窍，组合这些程序就会变得轻车熟路。更何况，运算程序一经写成便可以多次重复地使用。现在，我们对各个运算程序进行编号，运算0是相加，1为相减，2是相乘，以此类推。复制可以是运算5，移位是6，等等。随后，当我们再次使用寄存器机时，就完全可以用数字来发号施令了。

练习4（可选）

写出一个流程图及它的RAP程序，用来将寄存器机转变成一台便携式计算器。

a．寄存器2用来执行运算：

0＝相加

1＝相减

2＝相乘

3＝相除

b．在寄存器1和3中放入参与运算的数据。

那么“306”就表示“3 ＋ 6”，“513”表示“5 - 3”，“425”表示“4×5”，“933”表示“9÷3”。然后，我们用寄存器4、5、6、7来显示操作结果：寄存器4用于表示符号，其中0代表＋，1代表-；寄存器5显示答案的数值；寄存器6用于存放除法结果的余数；寄存器7表示警示，即对输入的内容进行报错，例如除数为0，或者寄存器2中出现了没有定义的操作的情况。

在这个例子中，寄存器中的内容（其中的数字）起到了四种不同的作用：表示一个数字，表示一种算数运算，表示数字的正负号，还有充当报错旗帜。

秘密2

寄存器中的数字会代表什么取决于我们所编写的是什么样的程序。

利用已经创造出的构造模块，我们可以建造出更厉害的操作程序。所以，只要有信心，我们就能写出一个程序并画出它们的流程图，例如：计算出寄存器7中数值的平方；或者是，算出从寄存器1到寄存器20这20个数字的平均数；或者，分解寄存器6中的整数，如果5是其中的一个因数，就在寄存器5中放入一个1；或者是比较寄存器3和4中数值的大小，将较大的内容放入到寄存器5中，但假如它正好大出一倍，就用寄存器7示警。

我们可以写出一个程序，它能检索100个寄存器，找出具有特定内容的那个，并把所在寄存器的地址写入寄存器101中。这是怎么做到的呢？先将那个特定的数字存入寄存器102，然后将它复制一份存入寄存器103。清零101，然后从寄存器1开始与寄存器103中的数字进行减法运算，被减数是寄存器103中的数字，但每次减量前都首先要在寄存器101中增量记录一次。需要找到的是减去特定数字后结果是0的那个寄存器。不是寄存器1就继而再检测寄存器2，以此类推，直到检验出相减结果为0的那个寄存器，这时减法运算停止，该寄存器地址也正好写入了寄存器101。在减量运算时，寄存器能及时“注意到”0的出现，正是这种基本的“感知力”使我们得以将寄存器机的“目光”转向它自身，让它也尝试检测自己的寄存器单元：来回移动其中的内容，在需要的地方切换操作，等等。

秘密3

寄存器中的数字可以表示任何事物，这说明，寄存器原则上也可以处理任何事物，它们可以“识别”所有用数字（包括用一个数字或是用一系列数字）表示的图案或特征。

举个例子，一个大型寄存器库就可以是一张黑白图片（任何黑白图片，例如这一页书的照片），每个寄存器都占一个像素，内容为0时显示为白点，为1时显示为黑点。现在，请书写这样一个程序：可以从成千上万的图片中搜索出一张白底上画有一条黑色水平直线的图片。先不要着急去做。生命很短暂，请先详细地考虑一下，要完成这项任务会遭遇到哪些困难和费事繁琐的步骤。通过想象，如果你真的设计出了水平线识别器、垂直线识别器或者半圆形识别器，那么，再请思考一下该如何将它们与另外几个有用的识别器融合起来去鉴别出一个大写的A，包括它成百上千种的字体。光学字符识别（OCR）软件可以通过扫描，把

页面上的图形非常准确地转换成一个计算机文本文件，这无疑是近代计算机程序的一次胜利。其中，每个字母或者数字符号都由一个美国信息交换标准码（ASCII）中的数字表示，所以我们可以搜索文本，还可以完成那些神奇的文字处理工作。利用的不是别的，就只是算术。那么OCR真的可以读懂些什么吗？不然。它并不理解摆在它面前的东西，它只是近似在读录。但对于我们那收纳丰富、装满了可动部件的工具箱来说，这无疑也是件不错的工具。

秘密4

既然一个数字可以表示任何事物，那么它就一定能表示一条指令或是一个地址。

寄存器中的数字可以表示指令，例如增量、减量、移动或者搜索；也可以表示地址，即寄存器在计算机里的地址。所以，我们可以用一连串的寄存器来存储一整套指令。现在，我们有一个主程序：程序A，它让机器通过一个一个地执行寄存器中的指令来完成要求的任务，然后我们可以在这些寄存器中存入另一个程序B。在我让机器运行程序A的时候，它首先要做的事情是去访问这些寄存器，而这些寄存器发出的是程序B的指令，机器随即执行。我想要指出的是，其实我们可以在寄存器机中央处理机中的一组寄存器中一劳永逸地保存下程序A，使其成为烙印在只读存储器上的“固件”，然后再利用它去运行程序B、C、D，等等。正是因为在寄存器机中设置了这个程序A，我们便把寄存器变成了一台存储程序计算机。

有了程序A，寄存器机才能忠实地执行我们通过数字输入寄存器的各项指令。寄存器机所运行的每一个程序都是由一连串数字组成的，它们顺次排列，程序A会依次访问它们，完成每个数字指定的任务。如果我们设置一套系统使这些指令互不混淆，比如，给每一条指令都以两位数字来命名（或任何相同位数的数字），那么我们就可以将构成程序B的一系列指令，例如，

86，92，84，29，08，50，28，54，90，28，54，90

改写成下面这一大串长数：

869284290850285490285490

这一长串数字既是程序B特有的“名称”，也是程序B本身，它需要通

过程序A一步一步完成执行。这里是另外一个程序：

28457029759028752907548927490275424850928428540423

另外还有：

890829647249028495249885674339043850388245980285454425

比较有趣的程序常常会有更长的名称，由成千上万个数字构成。你笔记本电脑里的那些程序，比如文字处理软件或者浏览器，就是由如此长串的数字写成的，通常由几百万个二进制数字构成。10兆大的一个程序写出来基本就是800万个0和1构成的一长串数字。

秘密5

所有可行的程序都能由一个单独的数字指代，都可以被看作是一串指令，等待通用万能机去执行。

杰出的理论家、哲学家阿兰·图灵想出的是另一种简单的假想机。它在被分成一个一个小方格的纸带上来来回回、嘎嘎作响，根据磁头在当前小方格中读到的数字，即0或者1来决定自己的动作（啊哈，这就是条件分支）。图灵机的工作就是轻轻地敲出比特（清除0把它改写为1，或者相反）或者保持比特不变，然后将纸带向左或者向右移动一个方格，进入下一指令。你可能会认为，仅靠一次移动一小方格，用二进制数字0和1而不是自然数0、1、2、3、4、5等来编写图灵机相加、相减程序或者运行其他功能，这听起来，比我们之前寄存器机的练习还要让人头疼，但事实上，两者的运行原理完全一致。带有程序A的图灵机就是通用图灵机，它可以从纸带上“读取”程序B，并利用纸带上任意位置的数据或输入来执行程序B。我们可以把王浩的寄存器机执行程序的过程分解为数字演算和按条件跳转，图灵的图灵机也是如此。两种机器都有着很强的提取程序数字名称并执行它的能力。我们要建造的是一个自带程序A的多用途通用机，而不是建造成千上万个固定不同功能的计算机器，让它们分别去执行各自复杂的任务。对于这台通用机，我们只要喂食各种程序（软件）它就能按要求办事，这就是虚拟计算机。

换句话说，通用图灵机就是一部通用拟态机。尽管名气稍逊，我们的通用寄存器机也是如此。你的笔记本电脑也如此。笔记本电脑能做的事情通用寄存器机也一样能做，反之亦然。但是先别大惊小怪，我并没有说两者有相同的运行速度。我们已经看出，寄存器机在处理某些问题

时极其缓慢，例如在进行除法运算时，它需要很费劲地一遍一遍地做减法，我的天哪！就没有什么方法来提速吗？当然有。事实上，自图灵机出现以来计算机的整个发展进程，专家们所做的一切就是要让寄存器机能做的一切工作变得高效起来，其他别无所求。

秘密6

自图灵的假想纸带机之后，计算机所有方面的改善，提高的都只是运算速度。

例如，冯·诺依曼为第一台严格意义上的可操作计算机所设计的那种架构就是为了提高机器的运算速度。他拓宽了交流窗口，让图灵机的磁头从一次只读取1比特变成了一次读取多比特。早先的一些计算机可以读取8比特或者16比特的“字”，有时甚至还可以读取12比特的字。而今天，一次读取32比特已经司空见惯，虽然这也成为了一个瓶颈（即所谓的冯·诺依曼瓶颈），但与当年的图灵机瓶颈相比，它还是宽多了。简单说来，字会从存储器复制到特定的寄存器（指令寄存器）中，一次复制一个，在那里再接受读录和执行。每个字通常由两部分构成：操作码和地址。前者如相加、相乘、移动、比较、遇零则跳转，后者则告诉计算机所要操作的内容在哪一个寄存器。所以，10101110　11101010101就表示，对地址为11101010101寄存器中的内容执行10101110操作，运算结果会显示在一个叫作累加器的特殊寄存器里。

寄存器机与冯·诺依曼机之间存在着一个巨大的差异：在寄存器机上，你可以在它的任何一个寄存器中进行操作，当然，只能进行增量和减量运算；而在冯·诺依曼机上，所有的算术操作都是在累加器中实现的，寄存器只负责存放复制、移动或者储存过来的内容以构成内存。硬连线可以单独完成各自的基本操作，这能省掉很多额外的移动和复制步骤。也就是说，有专门的电子路线负责加法，也有专门的电子路线负责减法，还有的负责遇零则跳转等等。操作码很像是区号或者邮编：指令要送到正确的地方才能执行。软件和硬件就这样相遇了。

今天的实体计算机中会有多少种原始操作呢？几百或者上千个？或者回归了旧时的美好，成了一台精简指令集计算机（Reduced Instruction Set Computer，即RISC），运作的几乎全是原始操作，但条件是，它们必须得以闪电般的速度运行。要想单纯用增量和减量指令去替代相加硬连线，并在速度上领先，条件是：对于运行步骤小于一百万

步的加法运算，增量和减量指令的运行速度要比相加硬连线的运行速度快一百万倍。

那今天的实体计算机中有多少个寄存器呢？数百万甚至数十亿个，但它们每一个容量是有限的，因此那些巨大的数字需要分散储存到许多寄存器中。一个字节是8比特，如果你电脑上有64兆字节的随机存取存储器（RAM），那么你就会有1 600万个32比特的寄存器或者是其他规格的等量物。我们已经知道，寄存器中的数字能代表的并不止是正整数。像π、还有1/3这样的实数是由“浮点”系统来保存的，它将数字拆分为底数和指数两部分，就像科学记数法那样（“1.495×10^{41}”），使计算机除了运算自然数以外还可以近似运算其他数值。浮点运算是以浮点数作为数值的算术运算，突出运用在乘法和除法中。20年前（这一章首版时），我们能买到的最快的超级计算机每秒钟可以完成超过400万次的浮点运算。

如果你还想要更快的速度，可以尝试将多台机器并行的方法，让它们在同一时间一起工作，而不是一个做完再交代给下一个做的串行式。并行机能做的事情与串行机大致相同，只是在速度上，它会略胜一筹。事实上，在过去的20年中，专家们一直都在致力于研究的并行机，大多是在标准的非并行冯·诺依曼机上实现的虚拟机。专用并行硬件开发出来以后，计算机专家们一面要忙着测算拓宽冯·诺依曼瓶颈所需的成本与最终获益，一面又要想方设法地提高数据传输速度，不惜调用协同处理器、高速缓冲存储器以及各种各样其他的途径。今天，日本富士通公司的超级计算机K-computer已经可以运行实现每秒超过一亿亿次的浮点运算。

这个速度基本上要与你大脑的即时运算速度旗鼓相当了。你的大脑是一台最卓越的并行处理机，包含着大约一千亿个神经元，每个神经元都相当于一个复杂的微型行动者，有着各自的行动规划。视觉“神经”有数百万神经元宽度的通道，单凭自己就可以把眼睛看到的视觉信息传入到你的大脑里。但神经元运行起来比计算机电路慢得多了，一个神经元在几毫秒（一毫秒是千分之一秒，不是百万分之一，更不是十亿分之一）之内可以转换状态并发送一次脉冲（可能是它们的增量或减量指令），而计算机使二进制数字移动的速度几乎接近光速。所以说，让机器更快速运转的一个至关重要的方法是要将它们做得更小。光走一尺大致需要十亿分之一秒的时间，如果你想让两个程序之间的通讯比之前更快，那就让它们离得更近些吧。

秘密7

没有再多的秘密了！

也许计算机最美妙的地方就是，它由各部分（操作）简单地组合而成，而各部分本身也很简单地组合，简单到没有什么地方是无法给出解释的。这里没有仙气，没有“形态共振”（morphic resonance），没有无形的力场，没有未知的物理定律，也没有那些神奇的组织。要知道，即使你用计算机成功地模拟出一些情景，完成这一切的也只是一些算术运算而已。

目前流行的量子计算机是什么呢？量子计算机能做普通计算机不能做的事情吗？算是，但也不完全。得益于微妙而奇特的“量子叠加”态，那些不被观测的物质能够同时处于“所有可能的”量子态上，直到观测导致“波包塌缩”。这个性质使量子计算机可以同时计算大量数值，解决诸多问题。（要想了解更多，请查阅你最喜爱的那些物理科普书籍或网站。）基本上，在计算机的提速进程中，量子计算机的发明是最新一次的重大变革，可以说，在提高计算机的处理速度上，它实现了一次巨大的突破。图灵机沿着它的纸带喳喳作响，寄存器机奔波于各个寄存器之间不停地增量减量，但时间紧迫，在几分钟、几个小时或者几天内，它们能做的工作也总是显得那么有限。日本富士通公司的超级计算机K-computer数万亿倍地加快了运行速度，但还是不能解决所有的问题，这在密码学中体现得尤为突出。量子计算机本该弥补这部分的缺陷，只可惜，我们还没有克服工程上的诸多难题，没有让它变得更加稳定、更加实用。就目前来看，每秒钟运行一千万亿次的浮点计算就已经是天方夜谭了。

25

虚拟机
模仿硬件运行的计算机程序

一般来说，实体机器由可活动的物质部件构成。设计者们常常会根据其功能去命名它们，如：割草机、开瓶器、咖啡研磨机。它们的形状设计和构造各不相同，利用的是不同的物理定律。但它们也有共同之处：拥有相同名字的机器都在某种程度上做着与名字的描述相差无几的工作。它们基本上都能满足使用者的需求，可能有些会稍微好用一点儿。有私家庭院的人会挑选一款噪音相对较小的割草机，小咖啡馆的老板则更需要那种能精确研磨出不同粗细粉粒的咖啡机，尽管它操作起来非常复杂。有些机器是多功能的，比如一把电钻，插上不同的附件就能变成一把电锯或是一部打磨器。

计算机与这个类似，不同的是它不仅能按要求做几件不同的事，它甚至能做成千上万件事。而且，它也不需要插上不同的附件，你只要点开不同的程序，即点开一串串由0和1组成的超长数字，就能让计算机内部实现一些必要的切换，从而转换成不同的装置去完成不同的工作。每套装置系统都是一台单独的小机器，没有齿轮、轴承、电线或者滑轮，这就是模拟机，“构成它们的是指令”。不像和面团、压纸浆或者造钢坯，计算机处理的只有信息。

所以，对于计算机来说，指令就是那些齿轮和滑轮。所有的信息都要转换成二进制代码0和1，它们是计算机能够和需要“读录”的唯一代码。印在芯片上的电路会数以万亿计地分流0和1这些代码，逻辑门不断地开合，将信息流输送到这条或者那条电路，并以这种方法来下达指令。计算机上唯一“可移动的部件”是硬件上数百万的小空位，它们可以在1态和0态之间任意切换。至于计算机在某一时刻具体是一台具有何种功能的机器，则取决于数千个至数百万个微型元件的设置。

若一台实体机的各部分相互配合可以呈现许多不同的状态，那它就具备极强的可塑性，只需要在其上施加一套特殊的指令，就能把它变成

一部虚拟机。由于虚拟机处理的是信息，那么它就能够与实体计算机做相同的工作：后者中的“可移动部件”只是硬件中的状态变化，而前者可以直接操纵状态变化来代表那些“可移动部件”。这也不难理解：你可以在一张纸上用铅笔演算一个长除法，但如果足够熟练，你也可以“在头脑中”进行这种运算，即通过想象，在脑海中的纸张或者黑板上书写过程。两者都会收到很好的效果，原因是，在面对这类情况时，我们想要得到的答案只是一种信息。相比之下，想象出的一个火腿三明治是不能用来充饥的，你得真的去做一个。计算机十分擅长在“头脑中”处理信息事务，所以我们有时真的分不清自己在使用和操作的到底是一台处理专项信息的“专用”“硬连线”机，还是一部在多用途芯片上运行的虚拟机。

今天，几乎所有的升降机、空调，还包括你的汽车、电冰箱和电视机遥控器里，都有那种微小、廉价的计算机芯片，它们实际上就是一些多功能计算机，能运行你笔记本电脑里操作的所有程序。但我们却选择让它们倾其一生只运行刻录在只读存储器上的一个简单的程序，将它们所有杰出的能力都封锁在那一两个技能上，例如点火控制或者周期除霜，等等。原因是，比起去制作那种只负责专项工作的专用芯片，这样做的成本要低多了。

虚拟机的概念是计算机科学中出现的一个最有用的想象支架，在计算机界，这早已人尽皆知。现在，到了将它引入其他领域的时候了。因为要从广义上使用虚拟机这个词（我会在适当的地方说明原因），所以我们有必要先了解一下它最初的意思，也就是说它本身的意思。“虚拟机”是计算机科学家杰拉尔德·波佩克（Gerald　Popek）和罗伯特·戈德堡（Robert　Goldberg）在1974年引入的一个概念，它最初表示的是“实体机的一种独立有效的复制品”，即各种指令的一种复制品。实体机是实在的硬件，由硅芯片和电线之类的东西组成，我们称它为A。而虚拟机是运行在另一台实体机B上的计算机程序，它能完美地模拟硬件A：速度可能会慢点，毕竟它要首先保障自己硬件上的基本操作有效运行，才能组织起硬件A上的那些操作，但的确都是同样的程序。如果你写了一个在硬件A上运行的程序，那么硬件B在运行模拟硬件A的虚拟机时，该程序应该能在硬件B上畅通无阻地运行。

这是一个非常有用的方法。部分原因是，它显然可以让一些资源免于浪费：假设这里有很多只能在Mac上运行的软件，但是如果你没有Mac，它们就无法使用。这时，如果可以在PC上编写出一个模拟Mac的虚拟机，然后成功运行，那么你电脑上的这些Mac软件就都可以运行

了。你的电脑会“假装”认定它是真的Mac，而那些软件则根本就浑然不知！试想一下，一个人胳膊骨折了，然后打上了石膏。石膏可以让骨折处不再活动，而它的重量和形状会让受伤者身体其他部分的动作产生一些相应的调整。现在，有个小丑想要模仿这个手臂上打着石膏的人，如果他技艺足够精湛，他就一定会让自己的肢体行动做出一些相应的调整，“显而易见”的是：他胳膊上打着石膏呢！如此说来，那些运行软件和我们这些旁观者可能就真的看不出此时是一台在运行Mac虚拟机的PC还是一台真正的Mac。

但现实与上面的例子却多少有些出入。虽然人们已经开发出了PC上可运行的Mac虚拟机，但据我所知，这个软件并没有宣传的那么可靠。另一方面，Mac却已经有了一个简单易行又非常可靠的虚拟机，它可以运行PC的操作系统Windows，借用它，Mac持有者可以运行所有自己喜欢的Windows软件。目前，绝大多数编写出的程序都需要专门的操作系统而不是专门的硬件，而相同的硬件上可以运行不同的操作系统。所以，虚拟机的概念还需要继续扩展，还应该包括对操作系统的虚拟模仿。一种操作系统本身就是一部虚拟机，它能让一个程序在略有差异的硬件上同样有效地运行。但一个操作系统毕竟只是一种软件，并没有模拟任何实际存在的硬件，它只是按照实际要求创建出虚构的机器，遵守相关的规定、接收特定的输入等等。

Java虚拟机（Java Virtual Machine，即JVM）已成为当今最流行的一种虚拟机，这是我们扩展虚拟机概念的又一个原因。Java虚拟机不是对任何一种实体硬件机器的模仿，它只作为软件机而存在，比较像一个操作系统。可以说，是Java的出现让今天的因特网变得如此丰富多彩。有了它，你可以从网站上下载一些小程序（Java的应用小程序），让你能够填纵横字谜、玩数独、探索地图、放大照片、与世界各地的玩家一起参加探险游戏。当然，它也能处理很多“严肃”的事务。用Java语言编写程序的网站不需要在意浏览网站的用户手上拿的到底是PC、Mac还是运行Linux操作系统的计算机。因为Java程序只在Java虚拟机上运行，而PC、Mac以及Linux操作系统上都有各自专门的Java虚拟机，匹配的Java虚拟机会自动下载到你的电脑上，几秒钟便可以安装完成。然后，各种Java小程序就都能运行了，犹如魔法一般。可能你有时会注意到Java更新正下载到电脑上，也可能你注意不到。理想状态下，你不用记住自己电脑上安装的是哪一款Java虚拟机，随意浏览网站就可以了，有时网站上的Java小程序能顺利在你的电脑里运行，而不能运行时匹配的Java升级软件就会自动下载安装，以确保能够继续运行。

从我说的这些虚拟机的扩展意义出发，我们可以看到，几乎所有的计算机程序都是虚拟机，因为它们都是由一系列指令构成的软件，一旦开始运行，就将通用计算机转变成了专用计算机，而这种专用计算机本可以专门设计并建造成专门的硬件。“通用”计算机，即今天我们所说的通用图灵机理念的提出，是阿兰·图灵为科学、也是为20世纪中期以来的人类文明做出的一项最杰出的贡献。只需要在上面简单地安装和运行程序，我们就能把它转化成任意一种功能清晰的计算机！（为了防止你想跳过这一部分，我在24章中已详细介绍过了。）你不用再去组装其他的硬件计算机，有一台就足够了，剩下的事情就交给软件吧。图灵机时代的开启让我们实现了一个非凡的想法：只要有一个大块头的硬件，在里面安装无数可塑可调的“记忆”盒子或者寄存器，把需要的指令放进这些记忆盒子里，让它们运行，这样你就把这个大块头的硬件变成了你想要的任何一种计算机。

图灵机，或是笔记本电脑，只能一条一条地执行指令。我们在这个基础上扩展出了“并行”计算机的概念，它可以一次执行多条甚至数百万条指令。硬件中的任何一个地方，只要它们本身保持着一种状态，收到指令后会变成另一种状态或者不变，那这个地方就是寄存器。这种状态可以是你电脑中一个比特的0态或1态，但不限于两种状态。以这种状态为基础，任何寄存器系统都可以完成一些特定的基础操作。例如，通过让寄存器由一种状态变为另一种状态，或者用寄存器的某一种状态判定下一步的操作，寄存器系统就可以按照“计算一个函数”或者“运行一个程序”来配置自己的各个寄存器。所以任意硬件上都可以运行利用了上述基本步骤的虚拟机。一种技巧可以使用一次就可以使用第二次、第三次、更多次，在硬件上部署一部虚拟机，虚拟机中可以再部署一部虚拟机，在其中再部署一部虚拟机……

请考虑一个下国际象棋的程序，它由Common　Lisp这种高级计算机语言写成，在一台PC上运行的Windows7操作系统上运行。事实上，这台PC佯装成了一台Windows机器，这台Windows机器又佯装成了一台Lisp机，这台Lisp机又佯装成了一台国际象棋游戏机。对于那些谙熟计算机或者国际象棋的观察者来说，要是从最上面一层观察的话，这个程序的诸多细节还多少可以理解（“啊哈，这个子程序生成了所有移动“象”的合法走法，然后调用另一个子程序来评估每一步……”）。然而，要是让他们去看这个程序实际的机器代码，看着那一串串0和1被吃进硬件的指令寄存器，这倒十足成了一个让人抓狂的好办法。所以，还是让我们把注意力放在那些较高的层次上吧。

其实，无论在哪一层次上，我们都只能只见森林不见树木，较低级层次上的那些细节总是很容易忽略。计算机上虚拟机的层层嵌套和关乎心智的小人儿功能主义中小人儿的层层嵌套，两者之间的雷同并非巧合。虚拟机推动着人类的创造和理解，实质性地完成了一些至今都超越我们想象的任务，比如完成航线预定、下象棋、预测天气、做记录等等，这些惊人的成就鼓励着我们：如果对大脑也采用类似（仅仅是类似）的技巧来实施逆向工程，或许我们也能得到一些收获。

在面对“说法语的人的大脑之间的相似性”这个问题时，排除解剖学上那些显著的差异，也许利用虚拟机的例子能为我们展示出一种最好的解释：所有说法语的人都具备某种版本的法语虚拟机，它藏身于他们大脑数以亿计的寄存器中，是一套由细微的习性和连锁装置构成的系统。说英语的人的大脑里同样也有这么一套固定模式的系统，即英语虚拟机。当你对一个法国人说，“Donnez-moi le sel, s'il vous plait.（请把盐递给我）”你基本上就能预见到法语虚拟机会作出的反应，因为这种反应与你，这样一个说英语的人为自己脑中的英语虚拟机输入“Please pass me the salt.（请把盐递给我）”时所给出的反应是一致的。但现在的问题是，要怎样在人类的大脑中建立起一个法语虚拟机或者英语虚拟机？

人们在下棋或是在说法语时，大脑各个不同的层面上到底都发生了一些什么样的活动，到目前为止我们还不得而知。[42]毫无疑问，我们对此给不出什么精确的描绘，不像是编译器，这种程序可以将高层指令翻译成可在硬件上运行的代码，只要计算机程序员能在最高层面上设计出一个想法，它就有把握吐出一个可执行的程序来。这里出现了一个非常重要的概念验证：至少我们找到了一种方法，它能够清楚地道出“具有数万亿活动部件的计算机，它的那种高等能力到底是什么”，而无需动用神奇组织。

26

算法

有效解决问题的策略机制

我曾在《达尔文的危险观念》中提到过，达尔文的那些伟大观念想要引出的是：

> 地球上的生命经由这种或者那种算法，在上亿年的时间里谱出了一棵布满枝桠的生命之树。

那么，准确来说什么是算法？事实上，这个概念有很多相互抵触的含义，而我取的是最广义上的那种理解。以下内容是我从《达尔文的危险观念》中摘录出来的，略有修改。

是达尔文发现了算法的力量。算法是一种特定类型的形式步骤，只要可以“运行”或者开始具现化，它就能合乎逻辑地产生特定类型的结果。其实，即使是在达尔文时代，算法也并不十分新颖。很多我们熟悉的算数运算，例如进行一个长除法或者结算支票簿，都是算法。另外，像走出井字棋的完美一步，还有按字母顺序排列单词时使用的那些判定程序，它们也属于算法。数学和逻辑学界对算法的本质和威力的理论反思出现得则相对较晚，它们是20世纪发展的产物，让我们开始重新认识达尔文的发现，直接关系到了计算机的诞生，而计算机的出现也让我们对算法的威力有了更加深刻、更加生动的理解。

要查看算法的词源，我们得向前追溯：经由拉丁文的algorismi到早期英语的algorism，之后又错写为algorithm的现代形式。再往上，算法这一术语其实得自于波斯一位数学家Mûusâ al-Khowârizm的名字，他在9世纪写过一本关于算术步骤的书，到11世纪，巴斯的阿德拉德（Adelard of Bath）或是切斯特的罗伯特（Robert of Chester）把它译成了拉丁文。数个世纪以来，算法这个概念表示的一直都只是那种万无一失甚至略显“机械”的程序过程，20世纪30年代，阿兰·图灵、库尔特·

哥德尔还有阿隆佐·丘奇（Alonzo　Church）这一代人经过了一系列创始性的工作，才多多少少确立起了如今人们对这个概念的一些理解。对我们来说，算法的概念有这样三个至关重要的特征，其中每一个总结起来都有些难度：

> （1）底层中立：不管你用的是哪种符号体系，也不管你用的是铅笔还是钢笔，纸张或是羊皮卷，霓虹灯或者空中文字，它们都不会影响到你运算长除法的过程。运算程序依据的是逻辑结构，各种物质材料的因果力量也许能为运算保驾护航，但它们不直接参与运算。
>
> （2）潜在无头脑：过程的整体设计可以非常华丽，结果也可能十分耀眼，但算法的各个既成步骤，包括步骤之间的过渡却出奇地简单。有多简单？一个听话的白痴或者是一个简单的机械装置就能执行。教科书中这样说：算法就像是为新手厨师设计的烹饪说明。一本写给主厨的料理书上可能多是这样的措辞："放入葡萄酒适量，将鱼煮熟。"而转换成算法之后，它会变成这样："取一瓶标签上写有'干'字样的白葡萄酒，用螺旋开瓶器打开，把瓶中的酒倒入锅底至3厘米深，然后打开锅下的燃炉旋钮调到高火，……"整个过程分解得相当仔细，最后都落实到一个个简单得不能再简单的小步骤上。你只需要充当一部食谱阅读器就可以了，不需要做出什么明智的决策或是精准的判断，也不需要发挥直觉。
>
> （3）结果有保障：不管什么样的算法，只要不出意外，它就能每次都按照同样的步骤运作。一种算法就是一份简单的、不易有出入的食谱。

显而易见，计算机就是靠这些特点成就的。每一种计算机程序都是一套算法：根本的组成元素是一些简单的步骤，它们的运行过程惊人地可靠，而运行机制却十分简单。电子电路是计算机常用的选料，虽然电子总是在硅芯片上飞来飞去，但计算机的能力却与它们的因果特性完全无关，运算速度除外。同样的算法，用在使用光纤分流光子的设备上就可以完成得更快些，而通过人力用纸和笔来计算就会变得很慢很慢。

达尔文发现的不是一种算法，而是一大类相关的算法，只是他还没办法对这些算法做出明确的划分。

27

让电梯实现自动控制
机器取代人类的逻辑

在离开这段有关计算机的插曲之前，我还想再多介绍另外几个有用的概念——源代码、注释和目标代码，它们对于我们探索“大脑怎样产生意义”这个问题有很重要的作用。一般来说，比起直接去解决一个智力难题，先研究透彻一个简单纯粹的相关实例、抓住概念的一些实质对我们来说可能会更有帮助。在人工智能领域，这种简单实例叫作“玩具问题”。先解决掉玩具问题，顺势再让我们去处理现实中那些恼人的大问题吧！下面所讲的是一个有关升降机操作员如何被计算机芯片取代的故事，我是为了简单纯粹才编造了这个故事，但这一点儿也不影响它的现实性。

在我年轻的时候，那时候升降电梯里还需要有操作员，他们每天的工作就是跟随电梯上上下下，在需要的楼层停下机器，让乘客们出入。在那个年代里，他们操作的是一根很奇怪的手柄，手柄以顺时针或者逆时针方向转动，这样就能使电梯上升或者下降。能让电梯准确地停靠在某一楼层是需要技术的，那时的电梯总是停得要比楼层地面高出或者下陷三五厘米，所以在乘客们出入时，操作员不得不提醒大家要多注意脚下。先到哪一层，到了要说些什么，还有怎样开电梯的门等等，这些都是有规矩的，操作员们要接受训练，记住这些规矩，不断地练习操作，直到把这些规矩都彻底变成了他们的第二天性。规矩本身是发明者们在设计过程中经过多年的苦思冥想才制定出来的，期间还会进行大量的微调和改进。我的设定是这样的：这种电梯的设计已经基本固定，设计师们留下了一本详细的操作手册。无论什么人，只要按照手册正确操作就一定能成为优秀的电梯操作员。

现在，让我们来想象一下，要是操作员所有的工作都能由一个简单的计算机程序来完成，情况又会是怎样？事实上，随着各种自动机械的发明不断出现，纯技术性工作已经开始与人工操作发生分离，我们想象

的这种情况也早已逐步实现，但现在我们的设定是，电梯从人工操作到彻底地实现计算机系统控制是一步到位的。

我们猜想，电梯制造商们会召集一批软件设计师，也就是程序员，把操作员一直使用的操作手册交给他们："这里有我们想要的产品规格。请根据书中的操作规则设计一个电脑程序，我们要求它能实现最棒的人工操作员所有的操作。"程序员们开始仔细对照这本操作手册，并列出了一个包括所有操作的清单，以及在哪些情况下操作必须执行或者禁止执行。在这个过程中，程序员可以避免手册中出现的那些不规范的情况。举个例子，如果给电梯安装上传感器，那么它以后就会在需要的层面准确地停下来，再也不需要操作员们像以前那样要根据不同的情况说"请向上迈步"或"请向下迈步"，升降机每次只需要简单地播放一句录音"第××层到了，请注意脚下"即可。那么接下来，需要程序员们去做的是要用一种名为伪代码的语言写出这个程序的草稿。伪代码是介于人类日常语言和更严格、更系统的源代码语言之间的一种混合语言，一串伪代码写出来大概是这个样子："如果呼叫楼层＞所在楼层，那么上升，直到呼叫楼层＝所在楼层即停止；开门。等待……"

如果伪代码条理清晰，能够达到想要的操作效果，那么就可以把它转译成源代码了。源代码是一种更加规范严密、更加结构化的操作体系，而且，它还包含了各种术语的定义，例如变量、子程序等等。人类其实可以轻松地破译源代码，毕竟源代码也是我们写成的。一旦找到突破口，你就会发现，电梯操作手册中的那些规定和术语在源代码里表达得清清楚楚。而且，源代码下面的这两个特点还会让这个过程变得更加简单：一、变量和运算的名称直接显示它们自身的作用，例如呼叫楼层（callfloor），总重量（weightsum），报楼层（tellfloor），等等；二、就像我们在第24章中已经说过的，程序员们可以为源代码添加注释，这些附加说明能让其他人明白他们写这些代码的目的，弄清楚代码各个部分的功能。你在编写一个程序的时候也应该添加一些注释，防止自己忘记当时设计这串代码时的用意，在重新对程序进行纠错时，这些注释也会相当有用。

源代码必须严格按照语法编写，所有的构成元素和标点符号都要写在正确的位置。只有源代码书写格式完全正确，编译器才能读录，然后将它们以目标代码的形式转化成基础的操作序列，供实体机或是虚拟机执行。我们不能要求编译器从源代码中揣测程序员的意图，源代码必须让编译器准确知道该执行什么操作。不过，编译程序有能力利用多种途

径去执行操作，它还会根据具体情况选择最有效的一种。编译器也有好有坏，如果将同一个程序的源代码输入到两个不同的编译器中，它们得出的目标代码程序很可能就会以不同的速度在计算机上运行。

举个例子：你编写了一个下棋的程序，然后将它的源代码输入到两台不同的编译器中，得出了两个版本的操作程序。将这两个程序输入同一台电脑让它们对抗。你会发现，尽管这两个程序“按照同样的顺序做着同样的思考”（当然会这样，因为它们源自相同的源代码），但也会分出胜负，因为两者的速度不一样，快的一方使用的基础计算周期较少，在有效的时间内它能预想得更多、更周全。

回到我们的电梯上来。编译器编译出目标代码以后，就生成了一个可执行文件，文件名中通常有扩展名“.exe”。这个文件就可以在虚拟计算机上运行了。也许还需要几经调试，比如回到源代码进行微调，甚至重写，它才能最终成为“完成”版的程序。我们将这套目标代码“烧”到一块小芯片的只读存储器上，而这块小芯片上通常带有一个通用计算机及其之上的很多层虚拟机，然后，我们将这块芯片安装到电梯上。安装时芯片的输入端要与传感器连接，传感器能接受到按钮发送的信息和电梯测量到的乘客们的总重量等；芯片的输出端连接的是效果器，它可以指挥电机转动来开门闭门、升高或者降低箱体，更新显示器，播放录音等等。

再见了！机器已经彻底把真实的人类取代了，而不只是取代了一个我们塑造出来的小人儿，它们就跟人类操作员一模一样，遵守着同样的规则。真的是这样吗？果真如此？好吧，但事实并不如此。它们只是近似于遵守着同样的规则罢了。这是一种很经典的中间状况，它既不像人类那样，能够逐字逐句地把规则保存在大脑中，然后随时查阅那些规则从而指导自己行动；也不同于那些行星，总是能不紧不慢地“按照”设定好的轨道运行。我们人类也常常处于这种中间状况。有时，通过不断训练一些显性规则就会内化或者变成一种惯例，我们甚至都忘记了它们，不自觉地就把它们丢弃了。比如，一般来说英语单词中字母“i”都会位于字母“e”前，除了这两种情况：一种是，前面的字母为“c”；另一种是，在“neighbor”和“weigh”中，发音近似于“a”时。一些还处于调试阶段的规则也似乎是可以遵守的：例如英语的语法规则，它对语言学家们来说一直是一个挑战。时至今日，他们费尽周折也没能写出一本标准的规则手册让我们把英语说得更好，而事实上，几乎每个十岁大的母语是英语的孩子都能调试建立起一些相当不错的目标代码，供他们脑中的英语虚

拟机使用。[43]

在结束这一章之前，我还想提醒大家一点：植入到源代码中的那些注释可以帮助程序员们去理解软件中相互啮合的各个部分的意图，但是，在我们为了方便描绘人类大脑所创作出的那些硬件、固件和软件中，它们却并不出现。自然选择在我们的大脑中安装的那些功能结构就像是一些没有注释的代码，这些功能是有目的的，但由于代码并没有对结构的目的有所注释，大脑也就无从理解这些意图了。（在第40章中，我们会有更详细的讨论。）在发育和学习的过程中也存在着未加注释的调整或者是未说明意图的调整。就像是语言学家们那样，我们还在奋力实施着逆向工程来破译所有那些“规则”和“程序”，虽然它会比用逆向工程目标代码来复原源代码更有难度，但理论上的确可行。

第三部分小结

几百年来，“大脑是灵魂的所在”这种观点一直在我们心中根深蒂固，而到了20世纪中期，它却突然变得荒诞了起来！我们都知道，大脑由很多形状怪异的器官结构组成，分配在左右两边，早期的解剖学家们给它们起了一些很生动的名字，例如海马体（形状像海马）、杏仁核（形状像杏仁）、褶皱皮层（看起来像树皮），那它们究竟都有些什么功能呢？它们不能消化食物或者净化血液吧？大脑就是一个可以冷却血液的器官吗？就像亚里士多德认为的那样，相当于一个散热器？这些器官通过神经纤维连接，因此可以彼此沟通。笛卡尔认为，神经纤维很像那种一头挂着铃铛的金属线，你拨动一头，另一头的铃铛就会有所反应，但这能说明什么呢？一声铃响就能让我们对大脑产生心智的过程理解得更深刻？但还有谁有更好的设想吗？(44)

后来就出现了图灵机，它是在巴贝奇、帕斯卡和莱布尼兹等人创立的理论传统上建立起来的。他们认为，大脑可以由一些简单的部分组成，这些部分从根本上说都是机械性的，就像一个老鼠夹、一枚响铃、一套钥匙和锁、一个突触，它们可以以一种巧妙的方式组织起来彼此发生作用，自主地完成一些智能工作，没有人类的干涉，也不需要有幽灵来统领它们。它们可以计算。

在图灵机发明之前，“计算者（computer）”其实是一种职业，比如说，那时候工厂和政府部门会雇佣成千上万的人做计算者，为商业、航海业、枪炮制造业以及银行业做计算。也许阿兰·图灵是这样猜测的，大脑也可能是一名计算者，它通过奴隶般地履行大量的指令去完成信息处理，这些指令非常简单，就像增量和减量那样。当这些早期的认知科学理论家们，像阿兰·图灵、冯·诺依曼、控制论创始人诺伯特·维纳（Norbert Wiener）、信息论创始人克劳德·香农（Claude Shannon）等人明确地将这个观点表达出来时，所有人都恍然大悟、深表赞许，并扪心自问：为什么我们没有想到这一点？大脑首先从感觉器官吸收信息，然后通过某种计算方式对这些信息进行处理。经过多重细致的计算处理之后，大脑能从这些信息中提炼出一些有关意义的非常有价值的矿石，大脑把它们分类储存，用于指导人体的行动，而人体的行动反过来又可

以为大脑提供必要的能量和保护。

图灵的设想最具创新力的地方在于，他淘汰了一些笨拙的环节，这类东西在早期设想的信息操作过程中异常明显：节骨点上永远得有一个员工，或是一个译者、一个图书管理员。总之，必须得有一个理解器，用于领会所有输入的信号。图灵明白，从某种意义上来讲，这种情况不可避免：在完成了对不同信号的辨识之后，还是需要智能操作为下一步的走向做出决定。图灵尝试要将这一过程中的理解成分降到最低：他发明了条件分支跳转，这个设置能够通过识别（近似识别）1或者0，A或者B，x或者y来确定（近似确定）自己应该选择左边还是右边的路径。分支跳转和算法，这就是这个系统所需要的全部了，只要在虚拟机上叠加虚拟机，在其上再叠加虚拟机，一个个叠加下去，就能组装出具有各种识别水平的装置。

50年来，我们一直都在为这一愿景努力，但一路走来却发现了很多问题，它们很细小，解决起来也非常不容易。如果大脑可以比作计算机，那它肯定不是我们平时所用的那种计算机。我们需要提醒自己计算机的基本特性，这样我们就可以从更符合生物原理的角度提出大脑的可能架构，而不是那些在我们的刻板印象中存有的商业架构。

这段插曲的观点已经大体澄清，其中还加入了大量的细节分析。你可以把它当作一种思考工具，帮助你更好地解决接下来的问题：首先，进一步探索意义在大脑中或者在其他机器里是怎样生成的；还有，在没有优秀的程序员和杰出设计者的情况下，演化的力量如何设计出了这整套精妙的构架。在后面，我们还要去面对意识和自由意志这两个概念，它们最为狡猾，但之前所掌握的这些思考工具能助我们一臂之力。

第四部分

更多关于意义的思考工具

INTUITION PUMPS

AND OTHER

TOOLS FOR THINKING

28

红发人那事儿
大脑中的认知机制

有个还算诱人的想法：大脑携带的全部信息，包括信念、看法、记忆、策略等，被分解成了类似语句的片段，而这些片段则被分类存档以待再次唤起。我们已经发现，这个想法存在着一些问题。脑写入（brain-writing）不能简单地置入一个虚假信念，而人们可以共享某个信念，例如发生在伦敦的那起谋杀，犯不着共享用脑语写成的公式。但是，大脑里还有什么可以储存信息呢？我们人类可以“零打碎敲地”进行学习，所以肯定有一些让孤立事实逐个地、粗略地加起来的方法。

经济学家和其他一些人常喜欢说，你不可能仅仅做一件事情。做“一件事”通常有诸多后果。同样，说你只能学习一件事物也不一定对。但大体上说这还是可能的。在先前的章节中，你了解到有一种叫作普度鹿的哺乳动物。除非你合上书去查资料，否则，除了它们会照料幼崽、长有脊椎并且相对稀少等事实，你不大能确认任何有关普度鹿的东西，因为如果它们不是那么稀有，你之前就该听说过这种动物了。你是如何学到这些并不神秘：你只是读到一句话并相信了这句话。但是，动物或尚未学会说话的小孩能否从一些有趣的零星经验中获知某个单一事实呢，比如简单句所表达的事实？知识、信念或学习过程一定得分解为句子般大小的碎片这一想法大概只是拟人论的幻象。人们在一天当中会碰上许许多多的陈述句，有些是口语的，有些是书面语的，据此得以获悉形形色色的事实，同样也会误信一些谎言。其中一些我们放在图书馆或者档案馆储存，而另一些我们只存放在大脑里。我们很少逐字逐词地记住一个实实在在的句子，但当我们记下眼前的一个句子的要点时，它一定是在大脑中被存储成了某种类似句子的东西，例如，一段用脑语写成的公式。真的是这样吗？如果不是这样，还能怎样？

设想帕特说“迈克对红头发的人有点看法”。帕特大概是想说，迈克对红发人有一种刻板印象，这当然意味着迈克对红发人有些不敬，也影

响了他对红发人的期望和同他们之间的交往。迈克并不只是对红发人抱有成见，他对红发人还有相当独特的看法。帕特可能是正确的，比他自以为的更正确！我们可能发现迈克确实持有某种东西，它不是想法或见解，不是信念或想象，也不是其他任何我们常说的能提供意识体验的东西，而是大脑中的某种“次人”的认知机制，该机制总归涉及红发人。也就是说，只要话题是关于红发人的，该机制就会系统地发挥作用，调整迈克认知机制中的各种参数，使他不大能接受或承认对红发人优点的各种恭维，在面对红发人时也更易产生相对激进的行为，等等。这个涉及红发人的运作可能十分复杂，也可能相当简单。

“迈克对红发人有看法”绝对为我们提供了一些东西，它的意义无法否认，然而，其意义无法以一个让人信以为真的语句的形式表述出来，这种表述顶多只是一个助记标签。所以，这种东西不能称为信念，无论是明确还是含混的，例如：所有红发人都是F开头的单词（“F”可以替换为任何一种对迈克的态度的表达，只要它足够公允）。毫无疑问，迈克对红发人抱有某种态度，但那不是一种特定的命题态度。用哲学术语来说，无论我们如何千方百计地在上述表达式中塞入不可兼条款、限定词、概率算符或是对其内容做显式调节，它都无法被处理为以下形式：

迈克相信：对于所有x，如果x是红发人，那么……

哲学家以及其他理论家往往想将所有认知状态“还原”为可以由上述公式表达的信息负载状态（information-bearing states），并把它们称为信念或者欲望。虽然这种手段是提供心理草图的好办法，也就是意向立场，但是想把它们高度精确地表达出来的希望依然渺茫。若是乐意，我们可以说，各种各样的信念在系统中是隐式的。这意味着系统当前的运行“基于这一假设”，即世界上的红发人有着一些如此这般的特征。电脑程序员有时会在源代码中加入一段注释，注明这个系统基于一套确定的假设，他们知道自己不必为精确表达耗费心力，因为这些注释不过是为观察者设置的助记标签，并不是由电脑近似读取或理解的东西，即便对于我们这些观察者来说，注释也不是有关内容的“详细说明”，与化学家在描述分子结构时用到的化学式也不尽相同。

为某种次人的大脑结构给出一种意向立场的解释，类似于为几段代码加上注释；注释得当，它就为我们提供了一个启发性的标签，但这并不等于把某个在信息处理过程中用脑语表达的公式译成了英语或其他什么自然语言。有些哲学家没能看透这一点，他们生造出一些内在语句操

控机制（internal sentence-manipulating machinery）的奇妙世界，在这些奇妙的世界里，用选言式谓词（我看见一个男孩或女孩）或者没有逻辑结构的谓词（我看见一个孩子）来表达一个特定的大脑事件的内容，是截然不同的。

这个直觉泵的目的是什么呢？它只是一次尝试，试图表明：对于力挺思想语言的那番熟悉的反诘“还能是什么呢”，或许会有一个令人满意的答案，到时候就能让那些认为思想语言显然成立的人无话可说。我真希望自己能给出一套宏大的计算架构，大获全胜地展示出一个可行的替代方案，可惜我不行。至今还无人能行，甚至几乎无人尝试，因为很多人依然坚信，思想语言就像某人很多年前说过的那样，是“最后的救命稻草”。但是切记，在认知科学领域，至今还没有人发展出一套可行的思想语言模式，甚至也不曾为此努力过。这是一个异常困难的问题。[45]但愿某个开放的心灵能受到这个直觉泵的启发，解决这个问题。

29

彷徨的双币机、孪生地球以及巨型机器人

原初意向性和派生意向性存在明显的边界吗

我把“迈克对红发人的看法”以及源代码写在了同一段落，这似乎在鼓励读者无视基础问题中的裂隙，更确切地说，是我关于意向性讨论中的鸿沟：原初意向性（original intentionality）问题，这个术语是塞尔在1980年发明的。表面上看，他在原初意向性和派生意向性（derived intentionality）之间做出的明确区分总体上令人满意，甚至引人注目。原初意向性学说主张一些人造物，例如书籍、电影、电脑以及指示牌，可能从我们身上获得某种意向性，而我们所具有的原初（或内在）意向性则完全不是派生出来的。比如印在本页的文字关乎哲学，恰恰因为作者和读者拥有关于哲学的思想和信念，我们设法通过这些文字传达有关哲学的思想和信念，若没有我们这些文字使用者，它们根本什么都不是。

相比之下，思想和信念的意义独立于任何隐秘的使用者；它们展露了原初意向性，它们是人造物的所有派生意向性的终极源头。这些人造物不仅包括词句、书籍，同样也包括地图、电影、印刷品、符号、标志、图表以及其他技术表征，最重要的还有电脑。你写在小纸片上或记在手机上的购物清单为什么能帮你买东西呢？仅仅是因为你了解这些符号结构的使用，你赋予它们意义，你有购物的欲望并且确信超市就是该去的地方。与购物单相比，这些东西更为直接，也更加根本。亚里士多德说上帝是不动的原动者（unmoved mover），套用这一说法，我们便是无意的赋义者（unmeant meaner）。

我们完全同意塞尔的话，任何东西都不能单凭其物理形状或其他类似的属性获得内在意向性。请看下图的字母形状：

F R E E B E E R

假设，由于某种机缘巧合，这些形状出现在火星的岩壁上，无论地

球人多想把它读作“免费啤酒”，它（“本来”）也不会是这个意思。这个图案不会是对任何东西的表示，不管它乍看起来像什么。如果这世界上有某些事件和物件是关于其他东西的，它必定会从它所处的表征和解释的意向系统中获得意义，该系统的状态，如信念、欲望、大脑状态等总得是已经具有意向性的才行。

接下来的问题是，是否存在具有原初意向性的东西呢？乍一看很明显，必须存在某些具有原初意向性的东西，因为派生意向性必须得从某种东西中派生而来。具有原初意向性的候选者显然就是人类心灵了。尽管一些有名的哲学家在许多方面都极不赞同塞尔，如杰里·福多尔和索尔·克里普克（Saul Kripke），但是，在原初意向性问题上他们倒是意见一致。他们和许多志同道合的人认为，人类心灵或心灵状态具有原初意向性，就此而论，人类心灵完全不同于机器人控制系统。

坦率地说，他们全都错了。是的，我不是在开玩笑。由于原初意向性和派生意向性之分有着无可争辩的魅力，任何敢于冒犯它的尝试恐怕都会被“好心人”误解：“他不会当真说我们错了，他一定是在借此表达一些深奥的哲学观点，只不过他愚蠢而荒谬地给它们披上了挑衅的外衣罢了！”也许，我只有提供一个生动、清晰的有关派生意向性的事例，才能让人们相信我是认真的。下面我会证明，通过一番更为仔细的考察，人们所钟爱的这组对照，即派生意向性与作为原初意向性的人类心灵，将被消解掉。这是一项艰巨的任务，但我还是要尽力而为。我需要三个彼此联系的直觉泵来完成这一壮举。

1．彷徨的双币机

设想一种制式软饮料自动贩卖机，它在美国设计组装，并配有传感装置，可以接受25美分的硬币。我们可以把它叫作双币机（two-bitser）。[(46)]通常，当我们往双币机中投入一枚25美分的硬币时，它便会进入Q状态，“意为”（注意双引号，此处只是近似意为）：“现在，我感到或收到了一枚货真价实的25美分硬币。”这种双币机设计得相当精巧，但并非绝对安全，它们确实会“出错”。当投入金属片或外国硬币时，它们有时也会进入Q状态，而有时它们也会拒绝合法硬币，未能进入本该进入的Q状态。

毫无疑问，我们可以检测造成机器“错觉”背后的模式。同样，具有足够的相关物理学以及传感机制设计参数知识的人至少可以断定某些“误

认”的情况。那么就可以直接从物理规律推理出：不仅合法的25美分硬币能触发Q状态，其他某种K型物件也能触发Q状态，但是过重的J型物件或者磁性的L型物件不能触发Q状态。如此说来，K型物件将是合适的金属片，它会可靠地“骗过”传感器。看看我在这段中多么不厌其烦地使用了“近似算子”，这样一来，在我给你们双币机的规格时就可以使用意向立场了。当你试着重写本段而不使用意向立场时，你会惊叹它多么有效，近似算子可谓不可或缺。

假使K型物件在双币机的正常使用环境中变得愈加普遍，我们就会期望双币机的拥有者和设计者研究出更高级、更敏感的传感器，以便可靠地区分真正的25美分硬币和K型金属片。当然，到时候狡猾的造假者可能会改变K型金属片的外观，因此，需要探测传感器的进一步升级，但技术上的升级可能带来收益递减，因为天底下本来就没有什么不会出错的机制。与此同时，工程师和使用者都清楚，这种基础版的制式双币机也能将就对付，为防范那些微不足道的造假者而做出技术上的升级太不划算。

唯有该装置的设计者、制造者、拥有者以及使用者之间的共同意向才能使之成为一部硬币探测器，而不是什么金属片探测器或者硬币/金属片二选一探测器。只有在这些使用者及其意向的环境或背景之中，我们才能挑出Q状态的某些情形，将其视为“切实”，从而把另外一些情形视为“出错”。也只有在意向的背景之下，我们才会一开始就将这种装置称为双币机。

我想，到目前为止，我已让塞尔、福多尔、克里普克等人点头称是了：意向正是如此这般地同人造物关联；这是一个关于派生意向性教科书般的案例，其间种种让人一览无遗。因此，没人怀疑，某种双币机直接从美国的工厂下线，贴上“A型双币机”的标签，就可以安装在巴拿马的软饮料贩卖机上，并开始工作，接受或拒收巴拿马的法定货币巴波亚25分硬币。通过设计和加铸币印可以很容易地区分巴波亚硬币和美元硬币，但它们之间不能通过重量、厚度、直径或材质加以区分。

我可不是信口雌黄。我从稀有钱币飞鹰专柜的权威人士艾伯特·埃勒尔（Albert Erler）处获知，标准的自动贩卖机无法区分25美分硬币和铸造于1966—1984年间的巴波亚25分硬币。这没什么奇怪的，因为后者就是在美国铸币厂用美元硬币的储备造出来的。虽然与这个例子不甚相关，但我想补充一点，只当满足大家的好奇心：2011年，巴波亚币兑美元的汇率是0.98，所以今天的一枚巴波亚25分硬币的价值略小于一枚25

美分硬币。

这样一个被带至巴拿马的双币机也会正常地进入确定的物理状态，无论将一枚25美分硬币、某个K型物件还是一枚巴波亚硬币投入该机器，我们都会以其物理特征来确认Q状态，但现在，被视为出错的情形不同了。在新的环境中，25美分硬币就像K型物件一样，被视为引起机器出错、产生“错觉”和误认的金属片。反过来，在美国，巴波亚硬币就是一种金属片而已。

当双币机安装在巴拿马时，我们能不能说所谓的Q状态依旧发生了呢？该装置“接受”硬币时的物理状态依然发生了，但现在是不是应该说，我们将那种物理状态确认为“识别出”一种新状态QB呢？没错，我们该怎样说有相当大的自由，毕竟双币机只是人工制品，对其感知和误认、切实或非切实状态，也就是对其意向性的讨论，简言之，“只是隐喻”。双币机的内在状态，或随你怎么称呼，既不真正（原初地）意味着“现在收到一枚25美分硬币”，也不意味着“现在收到一枚巴波亚25分硬币”。塞尔、福多尔和克里普克等人会坚称，它没有真正意味着任何东西。它的内在状态仅仅近似于意味着某种东西，但是，对于我们这些原初意向性理论的欣赏者来说，这足以引发一些问题。让我们仔细地看一看。

双币机最初设计为一种用来检测25美分硬币的机器。这是它的“应有功能”（Millikan, 1984），毫不夸张地说，亦是其存在的理由。如果没有这种功能，没人会花时间把它造出来。鉴于此，一般说来，我们可以将这样一部装置主要地或恰当地刻画为双币机，它的功能是检测硬币，因此，相对于这一功能，我们不仅能确认它的切实状态，即正确地检测出硬币，也能确认它的错误状态。

但这不能阻止我们为一个新的用途而临时征用双币机。但只要是物理定律，任何新用途都可以被允许，例如，用作K型物件检测器、巴波亚硬币检测机、门挡，甚至一件致命的武器。对于它的新角色，也许会有短暂的混乱期或不确定性。当积累了多长的记录之后，某物就不再是双币机，而是变成了巴波亚硬币检测机（q巴波机）、门挡或致命武器呢？[47]一台双币机，经过10年的忠诚服役，刚刚变成了一部q巴波机，对于这部q巴波机的首次亮相，它的Q状态是否是对巴波亚硬币的正确检测呢？会不会有某种习惯性的怀旧错误导致它将一枚巴波亚硬币误认作一枚25美分硬币？

就像上面描述的，双币机与我们人类截然不同，它没有任何可成为

过往经验的记忆，哪怕是其过往“近似”经验的“近似”记忆。但是，如果你认为这有影响，我们也很容易为它提供这种记忆。开门见山地说，假设原初意义上的双币机装有一个计数器，服役十年后它的记录为1　435　792。设想，它运往巴拿马时计数器并未清零，所以它在巴拿马首次工作时计数器会跳到1　435　793。这个数字意味着它还没能转接到正确地确认巴波亚硬币的任务吗？别忘了，它近似错误地将该事件当成了Q事件——检测25美分硬币；它的用途就是检测25美分硬币。这一主题的变种以及并发问题会沿不同的方向驱动你的直觉泵吗？拧开直觉泵上的所有旋钮，看看你的直觉到底发生了什么。

可以肯定的是，若严格地就双币机自身来考虑，不考量它先前的历史，双币机并没有什么内在的东西，使其区别于真正的q巴波机，也就是巴拿马政府委托订制的机器。然而，鉴于它的前世今生，它首次进入我们所谓的Q状态时，它的功能、用途和意义就没有任何问题吗？这是一个进入Q状态（意为“接收美分硬币”）还是QB状态（意为“接收巴波亚硬币”）的问题吗？

我和有些人（Millikan,　1984）都认为，它在巴拿马的首秀被视为进入Q状态还是QB状态，完全取决于在其新的工作岗位上，我们是否选它来检测巴波亚硬币。例如，它被巴拿马百事可乐特许经营人选中。如果它被选中，那么，即便新的运营者忘了重置计数器，它的首次“感知”行为也会被当作一次作为q巴波机的正确辨认行为，因为这是它现在的用途。它的功能就是检测巴波亚硬币。另一方面，如果一台双币机被阴差阳错发往巴拿马，它的首次亮相将毫无意义，尽管用不了多长时间，那些为了新目的而征用它的相关部门就能意识到并认可它的作用，随后，它接下来的状态也就被视为QB。在它用于检测巴波亚硬币之前，无论它检测巴波亚硬币的能力多么出色，它的接收状态也不会意味着（以其人造的、派生的近似意味方式）“现在接收巴拿马的巴波亚硬币”。

想必塞尔和他的同事对我这么说也会心满意足，双币机毕竟只是一个人工制品。它不具有原初意向性，因此也没有我们要试图揭示的什么“深度”事实。当隐喻式或者拟人化地谈论一部装置的状态时，他们会说，这不过是一个怎么说恰当的实际问题。

现在，我们已经牢牢地抓住了派生意向性，让我们看看非派生的原初意向性，即我们的意向性，有什么不同。在这一点上，塞尔、福多尔、克里普克等人不仅不同意我，也不同意哲学家露丝·米利肯（Ruth

Millikan）、保罗·丘奇兰德（Paul Churchland）和帕特里夏·丘奇兰德（Prtricia Churchland）夫妇、认知科学家侯世达、马文·明斯基（Marvin Minsky），以及工作在人工智能领域的每一个人。历经30多年的聚讼，大家的情绪依旧高亢。我们在争什么呢？

2．孪生地球

设想有个人，就叫琼斯吧，他望向窗外，以为自己看到了一匹马。窗外可能有也可能没有马。但塞尔及其同事会说，“他处在自以为看到一匹马的心理状态”这一事实，根本就无关于解释。这是一个粗糙的事实，也是一个原初意向性的范例。接下来，我们来构造一个与双币机高度相似的思想实验，看看会发生什么。这是一个归谬法的实例。设想有一个和地球几乎一模一样的孪生地球，除了那里生活着犸，我们这里生活着马。[(48)]犸看上去像马，除非让训练有素的生物学家用DNA测试，否则，可以说，它们几乎在任何方面都与马别无二致。但是，犸并不是马，就像海豚并不是鱼。孪生地球人可以管犸叫“马”“horse”或者其他什么名字，我们仅需记住：孪生地球与地球极其相似，除了那里生活着犸。

设想我们将琼斯送到孪生地球——犸的家园，对此他毫不知情。我们可以给他下药，安排他在孪生地球的床榻上醒来。当我们把他带至一匹犸前，他会很自然地说或者想：“看哪！一匹马。”当他这么做时，要么他确实被激发，进入了“相信自己看到了一匹马”的状态，有了一种错误的、非切实的信念；要么他被那匹犸激发，生平第一次切实地进入了“相信自己看到了一匹犸”的状态。我们如何分辨他进入了哪一个状态呢？由犸唤起的信念是真还是假？如果一开始他错误地认为“我看到了一匹马”是错误的，那么当他生活在这些犸中，并与孪生地球人谈论犸时，他需要多长的时间才会调整“马”这个音在其语言中的意义呢（期间他并没有意识到这一点）？如果他在孪生地球上还生养了几个孩子，那么他们从老爸那里学来的词“马”是指马，还是指犸呢？他们从未看到过马，记住，那里只有犸。

这显然是一个诡谲、极端的例子，但它确实带来一个好问题：是什么决定了我们词语的意义？怎么决定的？历史是否总是对所有东西都具有重要意义？当前对词语的使用能否克服或推翻历史？在本例中，琼斯并未向我们提供独一无二的洞察；他甚至不知道自己已不再生活在地球上，因此，他大概会坚称“马”意指马。他的词语，即从他嘴里说出的

词，其意义来自他的感知信念，他知道他所相信的：他在看一匹马。这就是他为什么说“看哪！一匹马”的原因。他可能还会加上一句“这不是显而易见的吗”。假设让我们告诉他这次旅行，并告知马与犸之间微小但重要的差别，接下来他会说什么呢？或者说他应该说什么呢？更重要的是，是否有个好理由支持他，无论他说什么都是确切无疑的？难道他不正是做了我们每个人都会做的吗？无论是他还是我们都不具有理论化地阐释这一处境的专享信息。假设他说他的词“马”现在意指犸，而且当他看到一匹犸并称之为马的时候，他并没有犯任何错。入乡随俗嘛。

他能否只是靠着声明其言其意就把问题解决了？如果随后他忘记他所做过的声明，又该如何？有时我们也会这样做：“从今往后，我就用‘渣渣’来指盐啦！麻烦把渣渣递给我！”在科学理论化阐释的背景中，这种规定性定义是一种重要且已得到确认的做法，但它依赖于一个交往双方的合作共同体。如果琼斯具有原初意向性，想必在任何环境中，都应有一个关乎其词语意义的事实。但是，琼斯本人似乎无法参考他自己的原初意向性，他并不能比我们这些观察者做得更好。举例说来，设想我们骗琼斯说，他被我们送到了孪生地球；他相信了（在哲学家的直觉泵中，人就能这么轻信！），随后，如果他告诉我们，他的词“马”现在意指犸，这时他是对的吗？也许他应当说他不知道“马”现在意指什么。那么，由于我们知道自己也可能被送到了孪生地球，我们是不是也得承认，我们不知道“马”意指什么呢？

我们这些对原初意向性这一想法将信将疑的人已经有了关于这些问题的答案，但还需要第三个直觉泵使之更为清晰，以便同传统直觉来次难得的对战。所以，请读者们擦亮眼睛！我要设法诱使你们放弃你们珍视的直觉啦。

3．巨型机器人

假设你想过25世纪的生活。已知的唯一办法是将你的身体保存在某种冷冻装置中，你会静静地、缓慢地进入沉睡，愿意待多久就待多久。你会在冷冻舱进入睡眠，直到2401年被自动唤醒。

设计冷冻舱并不是你面临的唯一工程难题，你还必须保护冷冻舱并提供必要的能量用来制冷或满足其他所需，使其可以运行400年。你不能指望你的儿孙负责日常管理，因为他们很可能活不到2401年，同样你也不大可能指望未来的后裔对你的生活计划感兴趣，你甚至可能根本没

有后裔。因此，你必须设计出一个超级系统来保护你的冷冻舱并给它供应运行400年所需的能量。

有两种基本策略可供参考。第一种，你应该找一个你所能想到的最佳场所，在那里固定装置会很好地提供水、阳光和任何冷冻舱以及该超级系统本身所需之物。这样一套装置，或“设备”的一个主要缺点是，当有危害来临时它不能移动。比如，有人碰巧想在此处修一条高速公路。第二种策略更加复杂，但它回避了这一缺陷：为你的冷冻舱设计一个可移动的装置，配有所需的传感器和预警装置，使它可以避开可能的危害并寻找所需的能量。简言之，就是造一个巨型机器人并将冷冻舱安置其中。

显然，这两种基本策略均来自大自然，它们的区别大致相应于植物和动物的分野。第三种策略同样来自大自然：孢子或种子能够待在它坚硬的护甲里几乎无限期地存活下去，但这并不适合你，因为你的生命支持系统需要高能量，而孢子处在尽可能低能耗的惰性状态。由于动物的策略与我们的目的颇为相合，我们假定你决意造一个机器人安置你的冷冻舱。你应该试着将它设计为可以“自主决定”有助于实现你的最佳利益的行动。笨拙的行动或错误的转弯会使它不适合保护你到2401年，而这是它存在的唯一理由。毫无疑问，这是一个工程方面的老大难问题：这需要通过顶级的专门技术设计出一个“视觉”系统和其他一些“知觉”系统来保障运动。还有，在整个冷冻期间，你处于沉睡状态，因此你不能提供指导，并为它的策略做计划，所以你只能将其设计为可以自发制订计划以应对环境的变化。它必须“知道”如何“找到”并“识别出”可用的能量源、如何移动到更安全的区域，以及如何“预见”并“避开”危险。既然有这么多工作要做，而且还要做得快，那你最好想方设法精打细算：切勿赋予你的机器人超凡的识别能力，它只要能识别在这个世界上需要识别的东西就够了。

再次注意，我将所有意向词或“心智”词，例如“知觉”“找到”以及“预见”，都加上了双引号，用来表示这是一种特定的近似意向性：派生意向性。这种意向性完全取决于人类的目的。这是你制作的东西，无论它具有什么样的意向性，它都属于你——它的制作者。若去掉了双引号，我就会被指责为在鬼鬼祟祟地搞某种意识形态，因为工程师等人往往因循守旧，会不加双引号地使用这类言语谈论某个信息处理装置的规格，例如升降电梯控制器。我有意没有这么做。为论证方便起见，我承认，使用意向语言对人造物功能进行的任何描述或规定都是隐喻。同样需要注

意的是，像双币机一样，机器人的机制服从经济上的考虑：它需要“检测”或“识别”许多东西，但它的“识别器”并非完美无瑕。它可以出错，但什么算作错误归根到底取决于制作者的需要和欲求。假使制作者想造一个在“认不清”东西的世界里四处瞎撞的、滑稽的机器小丑，那么，这些“错误”中就会有一些是弄对了的，是小丑控制器的巨大胜利。

现在让我们回到直觉泵。由于你不能指望你的机器人是唯一执行该任务的机器人，你的任务将变得愈发困难。如果你的奇想流行开来：你的机器人很可能发现自己在与其他机器人，以及你的人类子孙争夺有限的能源供应、新鲜的水、润滑油等（见67章关于其他行动者的重要性的简短讨论）。那么，毫无疑问，你必须将机器人的控制系统设计得足够精密，允许它能计算与其他机器人合作或者结成互惠互利联盟的利益及风险。但是，再提醒一下，这种计算一定是“马马虎虎”的近似，是迫于时间压力的任意简化。

这一设计方案的结果将是一个能够表现出某种自我控制能力的机器人，因为你本人一旦进入睡眠状态，就不得不放弃对机器人的精细控制和实时操纵。因此，它能够评估当前状态与始终保护你的终极目标，并从中得到它自己的附加目标。这些次级目标可能要上百年的时间才能实现，尽管你在设计阶段尽了最大的努力，但其中一些目标可能仍不明智。也许，在被其他机器人“说服”去保护他人之后，你的机器人一行动就与你的目的背道而驰，甚至是自杀式的。

注意，在这个时候，即便机器人所有的意识状态和行为都源自你的目的，它们也会开始变得有点脱离你的目标了。因为你设计的机器人在某种程度上是“自主思考”的，它的“思考”可能超出你的预期。

考虑现实世界中一个非虚构的例子，比如一台国际象棋计算机，它可以击败它的创造者。我们说计算机正在“琢磨”车的行动，“决定”不去王车易位，这么说的唯一理由是因为人类设计者设计它做那种事。但是，由于设计者的目标是造出一台优秀的国际象棋计算机，所以，对于计算机的状态（派生性地）应当关于什么的许多决策，其实是在倒逼自己：既然国际象棋程序需要游戏规则和游戏状态的准确信息，就一定存在着诸多涉及每个象、每个兵的状态，还有在电脑的“后”吃掉了对手的“马”时对整个棋局进行评估的状态，等等。如果计算机的状态没有恰当地关联到棋盘上每一个兵的位置，无论设计者怎么做，这个状态都不会（派生性地）关涉存活在棋盘上的兵的数量。一旦确定了设计者的最大目标，比如制作国际象棋程序、巨型机器人或飓风模拟器，暴虐的大

自然就会接管过来，并决定什么可行，什么不可行，哪个系统中的哪些状态是错误的或不准确的。

诗人可以通过一首看上去关于马但其实是关于教授的诗来表达自己的观点，比如威廉·布莱克（William Blake）告诉我们“愤怒的老虎比听从教导的马更有智慧”，然而计算机工程师却不能将自己的意图强加给他们的造物。

接下来让我们稍加盘点。巨型机器人精神状态的模造（simulacrum）只是这样：它并非在真的做决策、看、思考、规划，它不过是好像在做决策、思考和规划罢了。我们应该停下来想想这种说法包含了什么。我们设想的机器人肯定比不起眼的双币机复杂得多；我们赋予它“规划”新的行动方案、“从过去的错误中学习”、“养成忠诚”以及与其竞争对手“沟通”等能力。此外，为了能让它“规划”“学习”和“沟通”，我们不得不为它提供复杂的控制结构，赋予它自我反思或自我监控的能力。换句话说，它像人一样，可以通达自己的内部状态，当它“决定”不“想”向我们“撒谎”时，能够“报告”“坦承”以及“评论”其内部状态“需要”输入什么。它对这些状态的含义有“意见”，无疑，我们应该严肃地将那些“意见”视为充足的证据（这大概是我们能轻松拿到的最好证据），来证明这些状态在隐喻的意义上“意味着”什么。

切记，它不过是件人工制品，没有原初意向性；我们正在考察的是其派生意向性，对于观察者而言，其派生意向性与我们这些“真正”的行动者的意向性同样不明显。双币机没有能力动摇我们的解释性判断，因为它无法提供确定无疑的“声明”，表明它不知道自己现在身处巴拿马，或者看到巴波亚硬币让它感到意外。

对于这一直觉泵，可以有若干回应，我们将对它们略加检视。不过，我首先想引出我们在做出初始假设时所持立场的最显著含义：不管多少神奇的人工智能妙法植入其中，人工制品都不会有什么派生意向性。若坚持这一观点，我们就不得不接受，我们自己的意向性和机器人的意向性完全一样，因为我讲的科幻故事并不新鲜，它是理查德·道金斯（Dawkins，1976）看法的另一种形式：我们和其他所有生物都是“生存机器”，其目的是延续我们自私的基因。我们是被造物，是历经亿万年设计而成的用来保障基因的生存机器，因为基因无法出于其利益迅速地、知情地行动。我们设想的自身利益和基因的“利益”很可能相悖，即便我们存在的目的就为了基因的“利益”。基因的保存是我们的原初存在理由，虽然我们学会了无视这个目标，并仰赖基因业已安放在我们身上的

智力、学习能力，设计出自己的至善（summum bonum）。因此，我们的意向性源自“自私”基因的意向性。它们是无意的赋义者，而非我们！

当然，在任何意义上，我们基因的意向性都不是内在的；基因的“意义”首先取决于蛋白质的合成、发育，也就是说，取决于进化而来的、由ACGT密码子构成的“字母表”系统。原初只是说它是许许多多进化而来的表征系统（representational system）中的第一个。后来的所有系统都有行动者，即意向系统，行动者的表征从其推进的目标中获得意向性，就像巨型机器人的意向性那样。(49)

虽然这一图景为我们自身的意向性从何而来提供了一个满意的答案，但它似乎让人有点尴尬：我们自身的意向性派生自基因实体，而基因的意向性仅仅是类意向性。字面上的东西如何依赖于隐喻层面的东西？此外，在我的科幻小说和道金斯的故事之间显然有许多不同：在我的故事里，我假定，制造巨型机器人所涉及的工程是一个有意识的、慎思的、有远见的过程。即使如道金斯所言，我们是设计过程的结果，在此过程中，我们的基因是主要的受益者，那也是一个完全不存在有意识、慎思、有远见的工程师的设计过程。不过，我们马上就会看到，这不是一个好的反对意见。

自然选择理论的美妙之处在于，它告诉我们如何在关于种种起源的解释中将智能造物主排除出去。自然选择过程理应对这种非凡的、机巧的设计负责。基因不是设计者，它们本身再愚蠢不过了，它们不能推理、表达或理解任何东西。基因不能设计自己，它们仅仅是设计过程中的受益者，你可能会说，它们只是客户。在我们的故事里，基因可比作非常愚蠢、非常有钱的客户，他们雇佣最优秀的工程师为其建造生存机器。若不是他们，工程师就无钱可用，而且他们的存活正是其造物的代价。是谁或什么东西做的设计？当然是大自然母亲，或更直白地说，是自然选择压力下漫长而缓慢的进化过程。

在我看来，进化过程最令人着迷之处在于，它以其不可思议的能力映射出人类心灵（智能设计者）的某些性质。第五部分我们将更多地讨论如何思考进化，眼下，我想澄清一下可接受的意义理论和进化论之间强有力的联系。自然选择既无前瞻性，亦无目的性，虽然这怎么强调都不为过，但我们也不应无视这一事实：在对许多微妙关系做出无数有识别力的“选择”“辨别”和“鉴赏”之际，自然选择过程已然表明它对基本原理有着近乎完美的敏感。更挑衅地说，当自然做出选择时，它可以出于

特定的理由来“挑选”特定的设计，却从不曾有意或无意地“表现出”选择或理由。自然“选择”心脏作为血液循环泵是出于其优越之处，并非出于它们跳动时迷人的节奏，尽管后者也可能是自然“挑选”其他东西时的理由之一。

巴拿马百事可乐的特许经营人可以因识别巴波亚硬币的功能选择双币机，把它用作一部巴波亚硬币检测器，同样的道理，进化也可以因为血液供氧的能力选择某个器官，将其作为肺。只有相对于这类设计“选择”或进化“认可”的目的，我们才能确认行为、动作、感知、信念或常识心理的其他范畴。(50)

我们是自然选择设计出的造物，这一想法既熟悉又引人入胜；但有些人可能会说，这种想法超出了严肃论争的范围。(51)然而，为什么不仅是创造论者、“智能设计”理论家，就连塞尔、福多尔及其拥护者也有点儿下意识地抵制它呢？

我觉得，这一想法有两个相当不明显的含义让一些人难以接受。首先，如果我们“仅仅”是造物，那么，我们这些拥有思想的思考者就没有任何特别的权威来确认我们内心深处的思想意味着什么，甚至不能确认它们是否在根本上无所意味。用不着改变任何内在性质，双币机就可以变成q巴波机；过去常常用来指一种物事的状态现在意味着另一物事。所以，原则上，如果仅仅是造物，如果我们自己的意向性因此是派生的而不是原初的，同样的事就会发生在我们身上。例如，琼斯没有权威可以来认定他是在考虑马还是犸。其次，如果我们是这样的造物，我们非但不能保证有权通达深层事实来确定我们思想的意义，而且根本就不存在所谓的深层事实。功能解释有时是显而易见的，比如，心脏显然是一台泵，眼睛显然是用来看东西的。但是，当它并非显而易见时，比如当我们试图理解大自然母亲的心智时，那里并不存在任何可供解释的文本。当有关功能的“真相”（the fact of the matter）有争议时，也就是说，当多个解释都讲得通时，那里就不存在任何真相。

30

彻底翻译与蒯因式填字游戏
不存在绝对正确的翻译

当两个同样优秀的功能解释冲突时，有主张认为，根本不存在深层事实可以解决这一冲突，哲学家蒯因（Quine， 1960）以其彻底翻译的不确定性原理巧妙地捍卫了这一主张，他得益于一个著名的直觉泵。设想，在太平洋中心发现一座孤岛，岛民使用一种岛上特有的语言。由于没有双语翻译的帮助，人类学家或语言学家不得不通过观察、与岛民进行互动式试错来理解这种语言，这一任务被蒯因称为“彻底翻译”。蒯因认为，原则上，如果有两位调查员担负起为这种异域语言制作翻译手册的任务，他们大致可以做出相当不同但却同样优秀的翻译手册，并会为岛民的话语赋予不同的含义，因此不存在事实上正确的翻译。对许多哲学家而言，这种想法似乎过于极端，不值得认真对待，所以，他们对此只是不屑一顾，并继续固执己见。下面介绍的这种思考工具可以让上述想法看起来合情合理，甚至显而易见。有两件事需要解释：（1）蒯因的说法如何在“原则上”可能为真；（2）就算该说法原则上成立，我们为什么几乎无法给出一个实例。

我常用下边这道英语填字谜题考我的学生。片刻之后，他们中的大多数人都宣布做了出来。你可以在阅读下文前试试。

1.	2.	3.	4.
5.			
6.			
7.			

横排：
1. 脏东西
5. 一种伟大的人类需求
6. 使平整
7. 某影星

竖列：
1. 一种水上交通工具
2. 我们常常想要的东西
3. 正上方
4. 美国的一个州（缩写）

你做出来了吗？如果做出来了，你得到的是哪个答案？这道题有两组很棒的答案，它们藏在本书的第76章，在答案揭晓前给你一个机会做出这两组答案。这道题尽管小巧，但我花费了好几个小时才把它设计出来，因为横排和竖列之间的相互影响必须满足多重约束条件，它强烈地限制了设计的可能性。如果你有所怀疑，可以试着构造一个更大、更好的填字谜题！（如果你做到了，烦请发给我一份。我在其他地方会用得着的。）

有人会问，“第一竖列的单词应该是什么？”这代表了某种不合时宜的实在论。这道题根本就不存在唯一的答案。我是有意这样设置的。比如说，我并没有先定下一组答案（历史上第一个或原初的答案，“因此”是真正的答案），凑成这道填字谜题，然后再找出另一组答案。相反，我平时就收集了一些意思相近的四字母单词，从中我取出几对一并制定了这两组答案。

因为定义的规范允许有一些弹性，所以我们能构造出这样的一道谜题。这两组答案中都包含勉强满足定义的单词，但是，哲学家称之为整体论的那种严丝合缝的周边条件，却把这些单词放入了两个完全稳定的结构。你觉得有多大概率存在着与这两组答案平等竞争的第三组答案呢？通常，密码学家会抱持这样的信念：你若找到谜题的一个解决方案，你就已经发现了这道谜题的唯一解决方案。只有在特殊情况下才可能存在两个解决方案，但类似这道填字谜题这样的情况表明，存在一个唯一的解决方案并不是形而上的必然，而仅仅是在强大的约束条件下极其可能的一种结果。

人远比填字游戏或计算机复杂。人类复杂的大脑里充满了神经调质，大脑又连在身体上，而身体与整个世界深深地交织在一起，其进化史以及个人历史内嵌于世界，相互贯通，远超一个内嵌于语言共同体的字谜游戏。所以，米利肯等人是对的：鉴于设计约束的本质，在极端情况下不可能存在相异的解决之道，给出两个完全不同、有着全局不确定性且不分轩轾的解释。彻底翻译的不确定性在实践中当然可以忽略不计。不过，该原则依然有效。我们没有彻底翻译的不确定性，不是因为脑袋里存有“真正的意义”（蒯因称之为意义的“博物馆神话”，这是他的主要攻击对象）这个形而上的事实。

现实世界中我们没有不确定性，是因为有很多个独立的约束条件需要满足。密码学家的座右铭向我们保证了，这只是一个微不足道的烦恼。当现实世界遭受不确定性的威胁时，解决问题、确定解读的始终是

更具“行为性”或“倾向性”的事实，而非某种神秘的“因果力量”或“内在语义性质”。意向解释几乎总是竭力达到某个单一的解释，但是，可以想象，在极端情况下双重解释通过了所有测试，于是就不存在任何深层事实可以确定哪个解释“正确”。事实确实能够确定解释，但始终是“浅层”事实在发挥作用。

31

语义引擎和句法引擎
大脑只是通过句法引擎模仿语义引擎

意义如何能产生影响？它似乎并没有某种物理性质可以引发什么，如温度、质量或化学成分。大脑从激荡于感觉器官的能量流中提取意义，改善身体的发展前景，而身体则用以安置大脑并为之提供能量。大脑产生对世界上各种重要东西的预期，用这种形式“制造未来”，从而恰当地引导身体。大脑是一个非常耗能的器官，若不能完成这份重要的工作，它们就活不下去了。换句话说，大脑应该是语义引擎。大脑由大量的分子片段构成，这些片段的相互作用服从严格的化学定律、物理定律，对分子形状和分子力做出响应，换句话说，大脑事实上只是句法引擎。

设想你跑去找工程师，要他们为你打造一部验钞机或者其他类似的赝品检测器，要求它能将所有的真钞叠成一沓，而把所有假币堆成另一沓。工程师说，这不可能。我们造出来的东西只能响应“句法”性质，也就是各种物理细节：厚度、形状、纸张的化学成分、油墨的颜色模式以及其他难以伪造的物理性质。他们只能基于这类“句法”性质造出一部运作良好但并非万无一失的验钞机。做成你要求的那样太过浪费，间接、有瑕疵的检测就足够了。

大脑任何一部分的结构都受到同样的限制。无论大脑做什么，无论其输入意味着什么，或仅仅是近似意味着什么，都是由物理化学力引起的。虽然大脑有生命力，由蛋白质而不是硅和金属构成，但是，不要臆想大脑能凭借它内部的神奇组织直接地检测到意义。物理学永远胜过意义。直接响应意义的语义引擎像永动机一样，在物理上是不可能的。那么，大脑如何完成指定给它的任务呢？——作为句法引擎，它具有可以跟踪或模仿语义引擎的能力。[52]但这可能吗？一些哲学家辩称，假使大脑运作的微观因果解释是完整的，没有任何未解的空白，那么，根本就没有意义发挥影响的余地。在第33章，我们会遇到一个直觉泵，通过展示语义性质，例如真理、意义以及指称在一些简单的因果过程中发挥着

不容抹杀的作用，来证明上述看法是错误的。但在转向那个略显复杂的直觉泵之前，我打算检视一个更为简单的模型，它有助于驱散对哲学家的直觉泵的怀疑，若一切顺利，还可以消除一些可能会干扰理解的疑虑。

32

沼泽人遇上母牛鲨
哲学家最喜爱的直觉泵

直觉泵理应有效利落地工作，抽取待寻找的直觉，继而返回待命状态。但直觉泵往往会引发激烈的辩驳、反辩驳，对思想有所调整、有所扩展。20世纪美国优秀的哲学家唐纳德·戴维森（Donald Davidson）曾对我说，他后悔发明了这个直觉泵，因为它怂恿了一种毫无节制且不具持久启发性的争吵。下面所述的是沼泽人直觉泵，也是哲学家最为钟意的直觉泵之一，尽管它不算是戴维森的得意之作：

> 设想一道闪电击中了沼泽里的一棵枯树，而我此时正好站在边上。我的身体被分解为各种元素，与此同时，完全出于巧合的是，那棵树中完全不同的分子变成了我的物理复制品。我的复制品，即沼泽人，和我的行动一模一样；出于天性，它离开了沼泽，碰上了我的朋友，它似乎认出了他们，说着英语同他们寒暄问候。它回到了我的家，似乎开始撰写有关彻底翻译的文章。没人能看出不同。
>
> 但有一处不同：我的复制品并没有认出我的朋友。它不能辨识任何东西，因为它本来就不能对任何东西有所认知。它无法知道我朋友的名字，尽管它看起来知道，也无法记得我的家在哪儿。它根本无法领会我所说的“家”的意义，例如，它发出的“家”这个声音并不是在某个背景中习得的，而只有这样一个背景才能在根本上给出“家”正确的意义，或无论任何意义。实际上，我不知道我的复制品怎么能通过发出各种声音来意指任何东西，或拥有任何思想。（Davidson, 1987, pp. 443-444）

当诸如孪生地球、沼泽人这种话题成为正儿八经的思考时，其他领域的学界中人，特别是科学领域的专家们，常常就哲学家这种不可思议的娱乐发难，对此，哲学家并非没有注意到。是科学家没文化、精妙莫测的哲学探索于其不过是对牛弹琴，还是哲学家错失了对实在的把握？

我还是不要说了吧。

这些古怪的例子有意保留了某现象未被重视的一个特征，将其他所有特征的影响降至最低，并由此来证明概念上的要点，使真正重要的东西凸显出来。孪生地球一例设置了最大程度的内在相似性，即你被发送至孪生地球，却没注意到这次大转移，以便你的直觉能将外部环境的影响可靠地显示出来。而沼泽人直觉泵在让未来倾向和内在状态保持不变的同时，将“历史”的影响降到了最低。因此，此类思想实验在设计上模仿了科学实验，通过让其他变量保持不变，试图将变量间重要的交互作用隔离出来。这类实验的一个困难是，由于它们是直觉泵，因变量就是直觉，所以在产生直觉的过程中，想象力的贡献比哲学家认识到的更难控制。我们将拆除一些吊杆托架，它们实际上妨碍了读者的想象力，扭曲了他们的直觉，并因此使思想实验的“结果”变得毫无效力。

此外，这类实验还存在深层次的困难。凭空想出个例子就来“证明”深层的概念要点不过是小儿把戏。假设一头母牛诞下了某种生物，它同鲨鱼分毫不差。你若去问生物学家，它是鲨鱼吗？善意的反应可能是，你在费力不讨喜地搞笑。或者，设想在室温下，一个恶魔微微一笑就能将水变成冰。恶魔之水是冰吗？这种傻乎乎的假想根本不值得得到回应。一些哲学家认为，微笑的恶魔、母牛鲨[53]、僵尸以及沼泽人在逻辑上都是可能的，尽管它们在自然法则上（因果上）不可能。他们认为这一点很重要，但我不这样看。大概，广泛撒下反事实性网的动机是，我们获得的答案会告诉我们论题的本质。但是，这世道谁会相信真正的本质呢？反正我不会。

下面，考虑一个会问及磁性物质的类似问题，我们注意到，对于磁性物质，有两个相互竞争的候选“真理制造者”，即典型性质或本质：（a）所有磁性物质都能吸引铁屑；（b）所有磁性物质都有某种内在结构，且称之为M阵列。老式的行为准则（a）会不会最终被新的内在结构准则（b）所取代呢？或者，后者不过是在还原意义上对前者做了解释？为了一探究竟，我们必须想着为科学家提出以下这些沼泽人式的问题。设想你发现了某种物质，它能够吸引铁屑但却不像标准磁性物质那样拥有M阵列。你会把它叫磁性物质吗？或者，设想你发现某种物质具有M阵列却不能吸引铁屑。你会将其称为磁性物质吗？物理学家可能会回应道，若真见到了这些假想的物质，他们会留意那些更为重要的事情，而不是思考管它们叫什么。他们整个的科学图景有赖于磁畴中原子偶极子的排列与铁吸引力之间的深层规律，打破这一规律在逻辑上是有

可能的，但他们对这种“事实”毫无兴趣。“结构”因素和“行为”因素之间真正的协变关系才是值得关注的。如果发现规律不适用了，物理学家就会相应地调整他们的科学理论，重新安置这些术语。

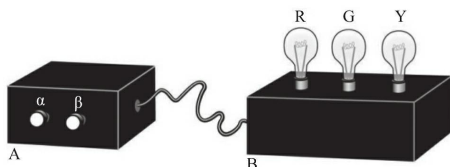
沼泽人是否有思想？是否说英语？母牛鲨是不是鲨鱼？它像鲨鱼一样游动，还能同其他鲨鱼成功交配。天哪，我没告诉你们吗？除了在它的所有细胞里都是牛的DNA之外，它同鲨鱼毫无二致。不可能吗？哲学家会说这在逻辑上也不是不可能。正是如此明显的不可能造成了后续没营养的讨论。正如戴维森的记忆“痕迹”出现在沼泽人的大脑结构中在物理上不可能一样，一条鲨鱼由包含牛的DNA的细胞构成同样在物理上不可能。沼泽人也许在逻辑上可能，如果仅仅因为想象某种宇宙巧合就能产生沼泽人这种东西，那么，根据定义，它们在逻辑上是可能的，但这种事从未发生过，谁会在乎如果它们发生了，我们该怎么说呢？

“我在乎啊，”注重反问的哲学家会说，“我认为，以最大的严密性定义你的用词以及所有逻辑可能的事项向来非常重要。这才是追寻真理之道。”是这样吗？在真实世界中，进化、发展以及学习的多重扭结将过去的历史和未来的作用束缚在一起。正是因为戴维森的身体有其特定的发展轨迹，假以时日，它就能形成戴维森所有的记忆、信念以及期望，这些自然累积的过程不会有一个真正的替代品。虚构的情形违背了这些限定条件，它所能达成的结论，就我所见，没有任何用处。事实上，这种费尽心力搞出来的例子给我的印象是，为某种假想的两分创造条件，好让你有恃无恐。“不，不，”哲学家说，“它不是虚假两分！为论证起见，我们悬置了物理定律。伽利略在其思想实验中忽略了摩擦力，他不也做了同样的事吗？”确实如此，不过两相对照，我们可以看到一个一般性的经验法则：思想实验的效用与其脱离现实的程度成反比。

孪生地球在物理上不可能，但其不可能并非沼泽人的不可能！千万别想用量子力学的多重宇宙解释，它虽表明孪生地球归根到底在物理上是可能的，但该理论只有在某些领域得到青睐；即便真的有无限多的宇宙，它们中又有无限多的星球与地球几乎一模一样，我们也不能发送地球人到这些星球访问。相比之下，双币机到巴拿马一游是可能的，有关它的种种事情也都会发生。不必悬置任何自然定律，我们就能设想我们希望设想的有关双币机的一切细节。

33

两个黑盒子
究竟是什么让红灯闪烁



有两个大型黑盒子A和B，它们之间由一条绝缘的铜电缆相连。A盒子上有两个按钮，标着α和β；B盒子上有三个灯泡，颜色为红（R）、绿（G）、黄（Y）。研究黑盒子行为的科学家观察到：按下A盒子上的α按钮，B盒子上的红灯会短暂闪烁；按下β按钮，绿灯会短暂闪烁。黄灯似乎从不闪烁。他们在各式各样的条件下进行了几十亿次的实验，并没有发现例外。于是，他们得出结论，认为其间存在着某种因果规律，并简单地概括如下：

α按钮导致红灯亮

β按钮导致绿灯亮

他们确信，这种因果规律以某种方式通过这条铜电缆传导，一旦切断电缆，B盒子就不会有任何反应；把两个盒子各自屏蔽起来，保持电缆连接，却不会破坏这种规律。他们当然很好奇，这种因果规律是如何通过电缆传导的呢？他们认为，也许按下α按钮会导致一个低压脉冲传至电缆，同时触发红灯亮；而按下β按钮将产生一个高压脉冲，它将触发绿灯亮。或者，按下α按钮引发单脉冲，触发红灯亮；按下β按钮则引发双脉冲，触发绿灯亮。显然，当科学家按下α按钮时，电缆中总是发生了什么事，而当他们按下β按钮，总是发生了一件与之不同的事情。揭示出这两个事情是什么，就能解释他们所发现的这种因果规律。

很快，对电缆进行的信号检测表明情况要复杂得多。无论何时按下A盒子上的按钮，一串脉冲流——开或关，或比特串，确切地说，10 000比特就会快速地沿电缆发往B盒子。但是，比特串的模式每次都不一样！

在比特串触发红灯或绿灯的情形中，一定有某种特征或性质。它会是什么呢？科学家们决定打开B盒子看看，当比特串送达其中时发生了什么。科学家在B盒子的内部发现一个普普通通的串行计算机，配有一个大型存储器，内含一个庞大的程序和数据库，当然也是由很多比特串写成。当他们跟踪比特串到达该程序时，并未发现什么奇怪的事情：输入串总是会以正常的方式进入中央处理器，在那里它会在极短的时间内引发几十亿次操作，并以1或0这两个输出信号中的一个结束，而输出1会打开红灯，输出0会打开绿灯。他们发现每一次都能在微观层面上解释这种因果规律的每一个步骤，没有任何困难或矛盾之处。他们并不认为运行过程有任何神秘原因，例如，当他们反复输入同一个序列的比特串，B盒子中的程序总是产生同样的输出：红灯亮或绿灯亮。

但这还是有点儿让人困惑，因为，尽管B盒子始终给出同样的输出，但其中间步骤并不相同。事实上，在产生同样的输出前，它几乎总是经历不同的物理状态。于其自身而言，这并不神秘，因为该程序存有每个输入信号的副本。因此，当同一输入信号一而再再而三地送达时，计算机存储器的状态在每一次会稍有不同。但是，输出总是相同的：如果某个特定比特串的首次输入使红灯变亮，那么从此以后，同样的比特串总是会使红灯变亮，同样的规律对绿串（科学家开始这样称呼这些比特串）也适用。科学家很容易就假定，所有比特串要么是引起红灯闪烁的红串，要么是引起绿灯闪烁的绿串。当然，他们并没有测试所有可能的比特串，而是仅仅测试了由A盒子发出的比特串。

科学家们决定验证他们的假设，他们临时断开A与B之间的联系，并略微修改了由A发出的输出信号。令他们感到迷惑和沮丧的是，一旦他们修改了来自A的比特串，黄灯就会闪烁！仿佛B盒子侦测到了他们所做的改变。但是，毫无疑问，B盒子能够接收人造的红串或绿串。只有当红串或绿串中一个或不止一个比特发生改变时，黄灯才会变亮，几乎总是如此。在看到一个“修改后的”红串变成了黄串，有人脱口喊道：“你们杀了它！”这还引发了一番猜测：红串和绿串在某种意义上是活的，也许分别代表了雌性和雄性，而黄串是死的。这个猜想很有吸引力，但它得不到什么结果，尽管科学家对数十亿有着随机差异的10 000位比特串进

行了实验，这些实验强烈地提示科学家，实际有三种比特串：红串、绿串和黄串，而且黄串的数量比红串和绿串的数量多了很多个数量级（更多内容见第35章）。几乎所有的比特串都是黄串。这让他们发现的红/绿规律更令人兴奋，也更让人费解。

是什么东西让红串打开了红灯，绿串打开了绿灯？毫无疑问，对于每个特定的情况来说，根本没有神秘性可言。科学家可以通过B中的超级计算机来跟踪每个特定比特串的因果效力，从而带着令人满意的决定论看到它在不同情况下点亮红灯、绿灯或黄灯。然而，单是检查一个新比特串，在没有“人工模拟”它对B盒子的影响的情况下，科学家无法预测它会产生这三种效果中的哪一个。他们从经验数据中得知，除非一个新比特串是由A盒子发出的已知比特串，否则一个新比特串成为黄串的概率非常高，几乎可以认为，任何新比特串都将是黄串，在这种情况下，一个新比特串只有十亿分之一的概率是红串或绿串，但是，不用B盒子里的程序来运行该比特串，就没人知道它究竟会是哪种。

也许，秘密在于A盒子。于是，科学家打开它，发现了另一台超级计算机，它有着不同的款式和型号，运行着不同的巨型程序，尽管如此，它也只是一台普普通通的数字计算机。科学家很快查明，这台计算机内部有一个“时钟”，滴答滴答地走着，每秒百万次，无论他们按下哪个按钮，计算机做的第一件事就是从时钟处获取“时间”，比如，101101010101010111，并将其分解成串，从而确定调用哪个队列里的哪些子程序，以及在其准备将比特串送至电缆期间首先访问哪一部分内存。

科学家弄清楚了，实际上，正是这个保持着良好随机性的时钟查询机制保证了同一个比特串从来不会被重复发送。尽管依赖于这种随机性或伪随机性，但它还是能够保证，每当科学家按下α按钮，计算机产生出红串；按下β按钮，送出的比特串最终是绿串。事实上，科学家们发现了几个异常情况：在大约十亿分之一的实验中，按下α按钮会发出一个绿串，或按下β按钮产生一个红串。这个小小的瑕疵吊足了科学家的胃口，他们想去解释这一现象。

后来有一天，两个负责制作这些盒子的人工智能黑客出现了，他们解释了这一切。（如果你想自己弄明白就不必往下看了。）制作A盒子的黑客Al，已经为某个“专家系统”工作了多年，该专家系统包含一个天底下所有“真命题”的数据库以及一部推理引擎，推理引擎可以从组成该数据库的公理中推断出进一步的结果。数据库中包含有美国职业棒球大

联盟的统计数据、气象记录、生物分类学、世界各国的历史，以及大量的琐碎数据。与此同时，制作B盒子的瑞典黑客Bo，为了打造自己的专家系统，一直在研发能与对手抗衡的“世界知识”数据库。在时间所能允许的条件下，他们都在往各自的数据库中塞入尽可能多的“真理”。[54]

但随着时间的推移，他们都觉得专家系统无聊乏味，该技术承诺的实际功能被大大高估了。事实上，该系统并没有很好地解决有意思的问题，也没能“思考”或“发现创造性的解决方案”。借着推理引擎，它们都善于以其各自的语言生产非常非常多的真语句，善于检测输入语句相对于它们的近似知识为真还是为假。因此，Al和Bo想看看他俩费尽心力搞出来的东西如何能够派上用场。于是，他们决定做一个哲学玩具。他们选择了一种通用语来翻译他们的系统，实际上，两个系统发送的都是以标准ASCII码写成的英语[55]，他们还把两部机器用电缆连起来。每当你按下α按钮，就会指示A随机或伪随机地择取它的某个“信念”，即某个存储在数据库中的公理或由公理集产生的推论，并将其翻译成英语（计算机中的英文字符已经是ASCII码的形式了），然后在末尾加入足够多的随机位使总位数达到10 000位，最后将该比特串发送至B，而B将输入的比特串翻译成自己的语言即瑞典版Lisp语言，并与自己的“信念”（它的数据库）做比对。由于两个数据库均由真理组成，而这些真理大致相同，依赖于它们的推理引擎，每当A发送一个它所“相信”的东西，B也同样“相信”，并以红灯闪烁来示意。而一旦A发给B一个它认为是虚假的东西，B会通过闪烁绿灯表示确实如此。

一旦有人篡改了传输过程，几乎总是会产生一个不合语法规则的英语语句的比特串，除非所有的修改只发生在末尾无意义的随机位。由于B对格式错误零容忍，所以它闪烁黄灯做出回应。假使有人随机选择了某个比特串，这个比特串就极有可能不是合乎语法的真理或谬误，所以黄灯总是占优。

所以，Al和Bo说，神秘的因果性质“红”实际上是英语真语句的性质，而“绿”是英语假语句的性质。突然间，让科学家多年来百思不得其解的探索成了小孩子的过家家。任何人都可以反反复复构造出红串。他只要写下“房子比花生大”“鲸不会飞”或“三乘四比二乘七少二”等语句的ASCII码就可以了。如果你想得到绿串，试试“九小于八”或“纽约是西班牙的首都”。哲学家很快就发现了一些小把戏：有些比特串在开始的头一百次是红串，尔后却变成了绿串。例如，ASCII码编制的这句话：“这句话送往你那里做评估的次数少于一百零一次。”

但是，一些哲学家说，比特串的性质是红或绿并不真就是英语语句的真或假。毕竟，有些通过ASCII码表达英语真语句需要数以百万计的比特位，此外，即便尽了最大的努力，Al和Bo在他们的程序中加入的也并非全是事实。例如，一些事以当时的常识看来为真，但在它们被收入数据库后，这些知识已被证明为假。为什么比特串的因果性质“红”并不等同于英语语句的真理性呢？理由很多。因此，也许这样定义“红”会更好：一段相对较短的表达式，用英语ASCII码写成，是近似信念几乎全为真的B盒子近似信以为真的东西。这个定义能让一些人满意，但出于种种原因，也有人吹毛求疵：它不够精确，或存在一些我们不能以任何非特例的形式排除的反例，等等。

但是，正如Al和Bo指出的，对于有待发现的性质，并不存在更好的描述，而且科学家念念不忘的不正是这种解释吗？难道红串和绿串的神秘性没有被完全消解掉？此外，既然神秘性被消解掉了，难道没有人发现，如果不使用某些语义学或心智学的术语，我们根本就看不到有任何解释我们故事开头讲的“α按钮导致红灯亮，β按钮导致绿灯亮”这种因果规律的希望？

一些哲学家辩称，虽然对电缆活动规律的新描述可以用来预测B盒子的行为，但说到底它不是因果规律。真理与谬误，以及刚刚考虑过的任何经过调整的替代者都是语义性质，其本身完全是抽象的，因此，它们不可能导致任何事情发生。其他人则反驳道，这是胡说八道，按下α按钮打开红灯正像是拧动点火钥匙发动汽车。如果发送到电缆上的东西无非是高电压、低电压，或单脉冲、双脉冲，人人都会同意这是一个典型的因果系统。该系统原本就是一部戈德堡式的机械(56)，但这并不意味着α和红灯闪烁之间联系的可靠性没有因果关系。事实上，科学家每一次都能描绘出确切的微观因果路径，并用来解释实验结果。(57)

其他一些哲学家被这番推理说服，他们开始论证道，这表明红、绿、黄这些性质根本就不是真正的语义性质或心智性质，它们模仿语义性质，仅仅是“好像”语义性质。真正说起来，红和绿是非常非常复杂的句法性质。但是，这些哲学家拒绝进一步说明它们是什么样的句法性质，也拒绝解释为何连小孩子都能快速可靠地举出关于它们的例子，或是识别出它们。尽管如此，这些哲学家坚信，必须对该规律做出纯粹句法层面的描述。毕竟，成问题的因果系统“只是”计算机，而计算机“只是”部句法引擎，不可能有任何真正的“语义性”。

“让我们设想，”Al和Bo反驳道，“如果你们在黑盒子里发现了我

们，做着同样的事情，那时你们的态度就会缓和，同意运作中的因果规律就是实情或信以为实情。你们能提出做出这种区分的充足理由吗？”该反驳迫使一些人表态：既然Al和Bo创建了各自的数据库，并将其作为自己信念的模型，那么，在一个重要的意义上，Al和Bo就已经在盒子里了。同样，这也使其他一些人否认说：这世界上不存在任何语义性质或心智性质。他们说，内容已经被消解掉了。尽管多年来一直众说纷纭，但我们一开始说的那种神秘性消失了。

堵住出口

黑盒子的故事到此为止。不过，经验告诉我们，天底下还没有哪个思想实验能清晰到让哲学家不误解它，因此，为了防止引发某些最具诱惑力的误解，我会不甚优雅地指出一些关键细节，并解释它们在这一直觉泵中的作用。

（1）A、B盒子中的装置不过是自动化的百科全书设备，它们甚至都不是“活的百科全书”，仅仅是“真理盒”。我们并没有在故事里预设或暗示这些设备是有意识、能思考的东西，或是行动者，除了在一个同样最低限度的意义上，我们可能会说温控器是行动者。它们是无聊透顶的意向系统，被严格地约束在一个单一的、简单的目标上。（当然，IBM的沃森同样如此。）它们包含大量的真命题以及某种推理机制，该机制可以产生出更多的真理，并能通过与数据库中已有内容对比来检测某个候选命题是不是“真理”。

（2）这两套系统是独立研制出来的，因此，很难认为它们恰恰包含着相同的真理，但出于戏剧效果以及我对它们在故事中所起作用的要求，必须假定它们之间有着非常大的重叠，这样一来，B就极有可能识别出由A产生的真理。我认为，以下两方面考虑可以让这一点变得合情合理：（i）Al和Bo可能生活在不同的国度，操不同的语言，但他们生活在同一个世界。（ii）尽管关于那个世界，即我们的世界的真命题的数量非常巨大，但Al和Bo都想打造出有用的数据库，这一事实将保证两套被独立研发出的系统高度重叠。虽然Al可能知道，在他20岁生日那天，他的左脚的位置更靠近北极点而不是南极点；而Bo也并没有忘记他的第一个法语老师叫杜邦，不过这些都不是他

们会放入各自数据库的真理。单单因为他们都热衷于打造各国可用的百科全书系统，就能保证他们各自的数据库有密切的对应关系吗？如果你对此有怀疑，那就添上一个不痛不痒的事实：在多年奋斗期间，两人就要讨论的主题交换了意见。

（3）为什么不是Al和另一个美国人Bob呢？或者，单就此事而言，为什么不干脆在B盒子里拷贝一个Al的系统呢？因为我不能让这种规律仅仅是显而易见的句法匹配，这一点对我的故事至关重要。在A进行语句生成任务、B进行语句翻译及真理检测任务期间，我们要避免被窥探到潜伏于数据查询结构之间的底层语义共性。这就是Bo的系统为什么采用瑞典版Lisp语言的原因。电子计算机作为一个物理系统，充其量也不过是一部句法引擎，它直接回应物理上可转换的差异，而不是意义。但是，A和B都被设计为一个尽可能接近我们想象中的无所不知的系统，一部语义引擎，其中装满了可被充分理解的真理。所以，当A和B这两种不同的句法系统被设计为如实反映同一个语义引擎的时候，为了解释它们所揭示出的那种引人瞩目的规律，唯一的方法便是上升到语义引擎的层面，在那里，真理得到认可，断言得到证实。我们的想法是创造出两个系统，它们的外部行为表现出令人着迷的规律，但内部则尽可能地不同，因此，只有当它们各自的内部结构同为一个共同世界的系统性表征时，才能解释那种规律。回想一下，这正是我在第13章想要的阐释的主题。

我们可能会停下来问，这样的系统是否真能如此不可思议，让逆向工程也对它无可奈何？或者换句话说，科学家们是不是真的困惑良久？鉴于密码学已经成为一个高度专门化、相当晦涩难懂的领域，一个人在回答上述问题前应当三思。我不知道是否有人能提供一个彻底有效的论证，来证明存在或不存在某种牢不可破的加密方案。但撇开加密术不提，黑客们都知道，当一段源代码程序编译完成时，所有源代码中的注释以及其他标注必须删除，只留下一段几乎无法破译的由机器指令组成的代码。“反编译”，也就是对目标代码进行逆向工程以及恢复源代码，有时在实践中是可能的（也许在原则上始终是可能的？），虽然它无法恢复注释内容，只是在高级语言中重新译出显著的结构。我认为科学家对程序进行反编译、解密数据库的努力会终成泡影。如果有必要的话，我们可以通过设想一种加密方案来强化我的观点。

在上述故事中，我们必须承认有件事非常奇怪：科学家们从来没有想过检查电缆中的比特流是不是经ASCII码翻译的。他们怎么能这么傻呢？有道理啊。你可以把整个装置，A、B盒子以及连接线，发送到“火星”来解决这个思想实验中的缺陷，让火星科学家试着找出其中的规律。他们同样也会看到α按钮导致红灯亮，β按钮导致绿灯亮，而随机比特串导致黄灯亮，就像我们看到的一样。但他们对ASCII码一无所知。除非火星科学家无意之中猜到每个盒子都包含着对某个世界的描述，而且它们描述的是同一个世界。否则，对他们来说，这份来自外太空的礼物将呈现出一种完全神秘的规律性，完全超出了所有的分析调查(58)。针对同样的东西，两个盒子担负着五花八门的语义关系，通过不同的“术语学”和公理化系统表达。这一事实才是这种规律性的基石。

丹尼·希利斯（Danny　Hillis）是连接机的发明者，连接机是20世纪80年代初他的公司“思考机器”（Thinking　Machine）率先建成的大规模并行计算机。当我在他那里试用这个思想实验时，他立即为这一难题想到了一个密码学的“解答”，然后还理所当然地认为，我的解答正好可以看作是他的解答的一个特例：“Al和Bo把世界当作‘一次性密码本’了啊！”这个词算是对标准加密技术的恰当说明吧。你可以通过想象另一种情形了解到这一点。比如，你和最要好的朋友即将被敌人（就算是太空海盗吧）抓住，他们懂英语，但对你们的世界知之甚少。而你们两人都懂摩尔斯电码，所以你们突发灵感想出了下边的加密方案：划表示讲真话，点表示讲假话。绑架者可以听到你们在说话。你说：“鸟会下蛋，青蛙会飞；芝加哥是一个城市，我的脚不是锡做的，美国职业棒球联赛8月开赛。”这样，无论你的朋友刚才问的是什么，你都在回答“No”（-·，---）。(59)而当下一回你需要说“No”时，你得用一些不同的句子。所以，即便抓住你们的人也知道摩尔斯电码，除非他们也能确定这些句子的真假，否则他们无法察觉是什么性质代表着点和划。

这种情形也可以作为噱头添加到我们的故事里，比如：我们不是把盒子里的计算机系统运到火星，而是把Al和Bo塞进盒子并将他们运送到火星。如果Al和Bo重施摩尔斯电码的恶作剧，那么，火星人也会对他们的行为感到困惑不解，就像他们对计算机的行为感到困惑不解一样，除非火星人得出结论：盒子里的这些东西需要语义解释。这对我们而言显而易见，但我们不是火星人。

这则故事的要点再简单不过了。没有什么东西可以代替意向立场，你要么采用意向立场的方式，找出语义层面的事实来解释这一模式，要

么永远被这种显而易见的因果规律所困扰。(60)

在这个节骨眼上，你可能像许多哲学家一样，又一次被这样的主张所迷惑：该直觉泵之所以产生这样的效果，仅仅是因为A和B两个盒子是人工制品，它们所具有的意向性纯粹是派生性的、人造的。它们内存中的数据结构获得的这些参照项（如果它们可以得到任何参照的话），间接依赖于感觉器官、生活史以及它们的创造者Al和Bo的目的。人工制品中的意义、真理或语义性的真正根源在于这些东西的设计者。Al和Bo具有原初意向性，而A和B都只有派生意向性。（当然，这就是为什么说在某种意义上Al和Bo都在各自的盒子里。）我其实可以把这个故事讲成另一个版本：盒子里面是两个机器人Al和Bo，在把它们塞进各自的盒子里前，它们“终其一生”在世界各地游荡以搜集事实。我选择了一个更简单的版本，是为了避免有人问及A盒子或B盒子是不是在“真正地思考”，但如果你想重新考虑这个思想实验以及与此相关的复杂情形，那么你要注意，巨型机器人的直觉泵已经对“真正的意向性不可能出现在任何人工制品上”这一想法提出了质疑。

第四部分小结

为了处理意义的基础概念，21种思考工具，包括十几个直觉泵和一些有用的概念悉数登场并投入使用。在它们的帮助下，我们做出了什么呢？它们是否全部奏效了？注意，证明一个直觉泵是否有价值有两种方法：首先，如果制作精良，它泵出的直觉就是可靠的、有说服力的，能够有力地遏制一些使人浮想联翩的错误路径；其次，泵出的直觉依旧可疑，但直觉泵可以帮助我们把注意力集中在它的前提上，看看出了什么问题。在这两种情况下，直觉泵都是一种杠杆，就像跷跷板，一边升，另一边就必须降，且唯有当它不会弯折或从中间断裂时方能奏效。直觉泵比跷跷板更复杂，因此，正如我在两个黑盒子的例子中所说的那样，我们需要拧一拧上面的旋钮，但毫无疑问，在接受它的结论前，我们还须检查其他旋钮。

对于意义，这一部分给出了什么样的结论呢？喜忧参半。意义不会是某种易于映射到大脑中的简单性质，我们也不会在任何地方找到可以解答某个句子、某个思想或某个信念的真正意义的"深层"事实。我们所能做的不过是，找到并锚定全部物理立场和设计立场以及现有资料的最佳解释。如果我们能找到解决蒯因式意义困境的一个方法，那么，几乎可以肯定，我们就已经发现了唯一的解决之道，这是说，我们可以确信，不存在某个尚未发现的更好的解决办法了。正如前往巴拿马的双币机表明的那样，意义总是相对于某个功能背景，它并不需要什么原初意向性，除了我们（近似）自私的基因的意向性。自私的基因从经自然选择而进化的功能背景中获得其近似意向性，而不是从一个智能设计者那里获得这种东西，就像那个订购巨型机器人的有钱客户一样。两个黑盒子的直觉泵表明，要理解并解释世界上的因果规律，意向立场，连同它接纳的所有近似信念之类的东西，并不是可选项。

在21世纪，我们对这些问题的研究大大增强了，第一次可以严格地、积极地对那些有着数以万亿计的活动部件的机制进行思考，它们的运作方式并不神秘。多亏了图灵，起码我们现在可以隐约看到一条路径，让我们从无理解能力的物质（物理立场）开始，经由一系列重新安排（设计立场，以及近似意义）达到这样一种自我理解：我们是信者、

知者、理解者的典范，经由意向立场简化为意向系统。

这些命题中的每一个都存有争议，或曾经存有争议，而且现在仍有许多专家不同意所有的这些命题。也许，将这些命题有序地组织起来会增强其说服力，同时我们会发现一个“临界质量”现象：它会吸引那些还不曾看到诸多部件是如何巧妙地啮合的人。另一方面，把这些命题像这样组织起来能让批评者更容易地找出一条贯穿其中的错误主线。不论怎样，我们都能由此获得进步。或许我们会发现其中的花招，它们遮蔽了真理，而不是照亮了真理。要找到其中的破绽，我们得回过头去转动更多的旋钮，看看会发生什么。最起码，发现这些破绽是一种辛酸的进步：它暴露出那些诱人的错误思想，而且几千年以来一直有哲学家身陷其中。

我并不是说该部分给出了一种意义理论。它给我们的只是一个相当宽敞的逻辑空间，如果我是对的话，一个恰当的有关意义的科学理论必须能容纳于其中。(61)既然我们对心灵的内容方面已经有了一个大致确定的了解，我们可否转向意识这一谜中之谜呢？还不行。我们需要建立更多的基础。理解意识的过程中浮现出很多问题，正如我们在这一部分已经看到的，这些问题都有着与进化相关的预设或影响。许多我们还在探索的主题已经涉及了进化，所以，在继续探索之前，让我们揭示这些主题，并对其加以澄清。这可是一个无限迷人的话题。

第五部分

关于进化论的思考工具

INTUITION PUMPS
AND OTHER
TOOLS FOR THINKING

达尔文认为万物经自然选择而进化，在我看来，这是人类有史以来最为杰出的想法，因为它大胆地尝试将物质与意义联结在一起，而现实的这两个方面似乎有着天壤之别：一方面，我们拥有心灵世界以及它们的意义，我们有目标、希望、渴求，以及最为尊贵同时也是最为平常的哲学话题——生命的意义。另一方面，我们头顶的银河不知疲倦地旋转，行星漫无目的地沿着轨道上升下落，无生命的化学机制

按照物理规律运行，这一切毫无目的、毫无理由。

然而，达尔文出现了，他告诉我们前者如何从后者中孕育而生，创造了意义，这是一种关于诞生的重要性的冒泡式（bubble-up）图景，它颠覆了传统的涓滴式（trickle-down）图景。自然选择的想法并不十分复杂，但非常强大，以至于让有些人无法直视，他们拼命转移注意力，仿佛它是一剂让人厌烦、难以下咽的苦药。这里有一些思考工具，它们有助于我们看到这个想法是如何照亮存在中阴暗的角落，如何将奥秘转变为我们可以解决的问题，并揭示出前所未有的自然荣光的。

34

万能酸
摧毁一切的达尔文思想

你可曾听说过万能酸？在学生时代，这一奇思妙想一度让我和同学们觉得很好玩。我不知道它是我们发明的，还是从其他人那里流传下来的，它和西班牙乌蝇、硝石一样[62]，都属于地下青年文化的一部分。万能酸是一种腐蚀性很强的液体，能腐蚀掉任何东西！但是，你怎么盛它？它能轻松地溶解掉玻璃瓶和不锈钢罐，就跟溶解纸袋一样。万一你不小心发现或造出点万能酸，会发生什么情况呢？整个星球都终将毁灭吗？它会留下什么呢？当所有东西与万能酸不期而遇并被其改变之后，世界会变成什么样子？我完全没料到，几年后我遇上了达尔文的想法，它同万能酸非常相像：它几乎蚀穿了每个传统概念，留下一幅发生了革命性剧变的世界图景，虽然大部分的旧地标还依稀可辨，但整体上却发生了根本性的改变。[63]

1995年，我首次引入了达尔文这一危险观念给出的图景，但许多因达尔文理论而感到恐惧的人也许是有意地错失了其中的要点。我费了些笔墨安抚读者，在万能酸席卷他们钟爱的伦理、艺术、文化、宗教、情绪乃至意识等话题之余，也留下了一些颇为精彩，甚至在许多方面更胜一筹的东西，只是发生了微妙的变化。达尔文的想法是一个革命性的想法，这点毋庸置疑，但它并没有毁掉我们珍视的那些东西，相反，它使那些东西的价值建立在更好的基础之上，并把它们与知识的其他部分优美地联结起来。几个世纪以来，人们认为，“人文与艺术”不但从科学中分化了出来，而且在一定程度上还免受科学审查的侵扰，但这种传统的隔离手段并不是保护我们珍爱之物的最好方式。将这些瑰宝掩藏于神秘面纱之下的企图会让我们无法在物理世界中为其锚定一个合适的位置。这是一种屡见不鲜的错误，特别是在哲学领域中。

一旦人们发觉自己钟爱的事物遭到了威胁，第一反应便是筑起一堵“坚不可摧”的高墙、一道马其诺防线。仅仅是为了多点安全感，他们

就决定在防御工事内部再圈出一些缓冲地带。这看上去是一个很好、很审慎的想法。似乎这样就可以防微杜渐，对那些可怕的灾难防患于未然，但人人皆知它们会得寸进尺。那就深挖洞！高筑墙！尽可能地做好准备。但这一策略往往让捍卫者为一大堆难以置信、站不住脚的教条所拖累，所以到最后，他们的捍卫一定会变成绝望的孤注一掷和声嘶力竭的无奈呼喊。

在哲学上，这一策略往往表现为某种绝对主义：人类生命的神圣性是无穷无尽的；神圣、不可思议的天才寓于伟大艺术的核心；意识问题太难了，我们这些肉身凡胎难以理解。还有我最喜欢攻击的靶子之一，我称之为歇斯底里现实主义（hysterical realism）：总是会有解决意义难题的深层事实。这些事实是真实的，是真正的真实，虽然我们无法系统性地发现它们。这确实是一个诱人的想法，部分原因在于它求助于凡夫俗子的谦卑感。我们凭什么说，没有任何事实可以解决这些问题呢？爱因斯坦对量子力学不确定性原理的抵触众所周知，他是这一诉求最可敬的示范。“上帝不掷骰子”是他的衷心之言，但这种抵触情绪根本没有合理根据。你平心静气地想想，我们或爱因斯坦凭什么说上帝不掷骰子？你会发现这不大可能是处理此类问题的方法。稍后，我们会看到歇斯底里现实主义反击的一些机会，并了解如何抵制它。进化论的思路是一剂解毒灵药。

35

孟德尔图书馆

只有极少数的A、C、G和T组合才是有意义的

人类基因组测序的工作已由克雷格·文特尔（Craig Venter）[64]等人完成，但这到底意味着什么呢？不是说每个人的DNA都不相同吗？是的，非常不同。事实上，单凭留在犯罪现场的一小段DNA，我们就有超过99%的把握确定凶手。然而，人类DNA也非常相似，只要给出其他物种DNA的片段，科学家就能将其与人类DNA区分开来。这怎么可能？我们每个人的DNA怎么会如此不同，却又如此相似呢？

理解这一惊人事实的一个好方法是将书本中的文字与DNA做一番对比，阿根廷作家豪尔赫·路易斯·博尔赫斯（Jorge Luis Borges）提出了一个小巧的寓言“巴别图书馆”（The Library of Babel, 1962），形象地说明了这种差异性和相似性的共存之道。博尔赫斯讲述了一群人进行着孤独的探索和猜测：他们发觉自己置身于一个巨大无比的图书馆之中，它的结构很像蜂巢，由成千上万甚至上亿个六边形通风井组成，这些通风井四周布满露台，露台上排放着一排排书架。站在露台的栏杆边上向下探视，没人能看到尽头，也从来没有人发现哪个通风井不被其他六个通风井包围着。于是人们怀疑，图书馆会不会是无限的？最终，他们认为它不是无限的，但也可能是，因为书架上似乎杂乱无章地躺着所有可能的书籍。唉！

假设每本书有500页，每一页包含40行，每一行有50个字符，那么每页就有2 000个字符。每个位置要么是空的，要么被印刷字符所占据，这些字符是从100个字符中选出的，包括英文以及其他欧洲语言的大小写字母，外加空格和标点符号。[65]巴别图书馆某处的一本书可能全部由空白页组成，而另外某处的一本书的内容则可能全是问号，但绝大多数的书的内容都是由乱码构成的，没有拼写规则，没有语法，也毫无意义。每本书有500乘以2 000，即100万个字符位，所以，如果我们用这些字符的排列去填满一本书，那么，巴别图书馆将有$100^{1\ 000\ 000}$本不同

的书。然而，据估计[66]，我们目前能观察到的宇宙区域中只有100^{40}个左右的粒子（质子、中子和电子），所以，巴别图书馆是一个物理上绝对不可能的对象，但由于博尔赫斯在他的想象中建构了严格的规则，我们倒可以清楚地想一想这件事。

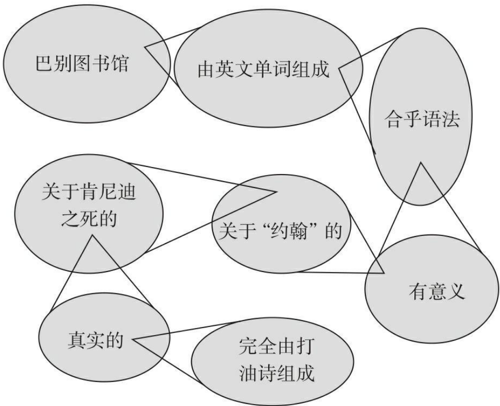
这会是所有可能书籍的真正集合吗？显然不是，因为它们被严格地限制“仅仅”用100个不同的字符印制而成，可以设想该字符集并不包括希腊语、俄语、阿拉伯语以及中文字符。因此，这个图书馆中肯定缺失了许多极为重要的真实书籍。当然，巴别图书馆藏有这些书的英文、法文、德文、意大利文等语言的精良译本，同样还有不计其数的低劣译本。那些超过500页的书会分册印制：从一册到另外一册，中间没有任何间断。

一想到在巴别图书馆中必定有某些书让人觉得很好玩，比如一本整整500页关于你的翔实传记，记录了从你诞生之日起，直到你逝去的那一刻之间所有的事情。然而，想要在图书馆中找到它几乎不可能，因为这里边还存有大量关于你的其他传记，准确记录着直到你10岁、20岁、30岁、40岁生日为止的所有事情，而这些书的后续部分则以相当不同、相当有趣的方式，彻头彻尾完全错误地记录了你的余生。即便如此，在这座庞然大馆中找出一本可读的书也还是希望渺茫。

我们需要一些术语来称呼这里涉及的数量。巴别图书馆并不是无限的，所以在其中找到任何有意思的东西的机会也并不是真的微乎其微。[67]这些话都是耳熟能详的夸夸其谈，但我们应该避免这种谈论方式。不幸的是，所有流行的有关数量的隐喻，九牛一毛啊、大海捞针啊、沧海一粟啊，都远远不够。甚至任何实际的天文数字，如宇宙中基本粒子的数量、宇宙大爆炸以来以纳秒为单位所记录的时间长度，在这些庞大但却有限的数字背景中都会湮没不见。如果在巴别图书馆中找出一本可读的书就像在大海中随便找出一滴水那般容易，那我们倒不妨一试！随意来到图书馆一隅，找到一本哪怕只包含一句合乎语法规则的句子的书的机会也是相当微渺，我们应该好好领会“微渺”（vanishingly small）这个词，同时，我们给它配了一个对偶词“浩瀚”（vastly），作为“远远超乎天文数字”（very-much-more-than-astronomically）这一意思的简称。[68]

这里还有另一种方式来理解巴别图书馆是如何大得离谱。正如刚才所述，在这些书籍当中，只有其“微渺”的子集是由英语单词组成的。该子集本身仍然异常“浩瀚”，只有“微渺”子集中的单词是按照语法规则排

列的，而数量“浩瀚”的书籍中充满了这样的字符串：“好是因为巴黎帮助容易来自民主的脱衣老虎”。在那些合乎语法的书籍当中，存在某个“浩瀚”但却“微渺”的子集，它由有意义的句子构成，其余书籍由合乎语法规则的句子构成，但这些句子随机选自那些合乎语法的书籍；在那些有意义的书籍当中，又存在某个“浩瀚”但却“微渺”的子集，是关于某个名叫约翰的人的书籍集合；而在这个子集当中，又有某个“浩瀚”但却“微渺”的子集，是关于约翰·肯尼迪遇刺的；这其中又只有某个“浩瀚”但却“微渺”的子集讲的是肯尼迪的真人真事；在这些真实的书籍当中，还有某个“浩瀚”但却“微渺”的子集完全由打油诗组成！没错，这些可能存在的关于肯尼迪之死的戏谑之作，比美国国会图书馆中的卷宗还要多！不过，这里边任何一本书基本上都未曾出版过，这倒也是一件好事。



巴别图书馆中藏有《白鲸》（*Moby-dick*），而且还收藏着与正版《白鲸》相差一个印刷错误的变异版《白鲸》，为数100 000 000部。这并不算一个特别“浩瀚”的数字，但一旦我们加入相差1个、10个甚或1 000个变异字符的变种，变异版《白鲸》的总数就会迅速升高。即使一本书里有1 000个变异字符，相当于平均每页两个变异字符，我们还是可以明白无误地将其识别为《白鲸》，而这些书的数量将是非常“浩瀚”的。不过，你读到的是哪一本根本无关紧要，你甚至可能一本都找不到！几乎所有的《白鲸》都同样精彩，并且都讲述了同一个故事，除了可以忽略不计的、几乎无法识别的差异。然而，它们也并非都差不多。有时候，一个关键位置的变异字符很可能是致命的。另一位颇具哲学趣好的作家彼得·德弗里斯（Peter De Vries）曾有部小说(69)这样开头：

“给我打电话，以实玛利。”（Call me, Ishmael.）(70)

哇，看一个逗号多有用！当然，句子也可能会变异为："把我团成球以实玛利。"（Ball me Ishmael.）

在博尔赫斯的故事里，所有的书都杂乱无章地摆放着，但是，即便它们严格按照字母表顺序排列，也无法找到我们想要找的书，例如《白鲸》的"精华"版。试想一下，我们乘飞船在巴别图书馆的《白鲸》星系畅游。由于这个星系本身大大超过了整个物理宇宙，所以就算你能以光速沿任意方向行进几百年，你所看到的东西也只是与《白鲸》几乎没有任何区别的副本。你永远不会看到其他任何东西。我们知道，通过一点点积累印刷错误，有无数多条路径可将一部伟大作品变成另一部作品，但即使我们知道一定存在一条最短路径，从《白鲸》到《大卫·科波菲尔》（*David Copperfield*）也是遥不可及、远到难以想象的。如果你发现自己在这条路径上，就算你手头上有这两本书的文本，你也会发现，单凭局部观察几乎不可能发现哪个方向能通往《大卫·科波菲尔》。

换句话说，这个逻辑空间太过"浩瀚"，因此我们无法直接应用许多关于定位、检索和查找的通常办法以及其他诸如此类的流行做法。博尔赫斯让书架上的书以随机顺序摆放，这一画龙点睛之笔为我们带来了一些可谓令人赏心悦目的反思，但让我们来看看如果他设法以字母表顺序安排所有的书，会为自己制造怎样的难题吧。由于只可用100个不同的字母字符，所以我们可以将其中的某一特定顺序当作字母表顺序，比如，a、A、b、B、c、C……z、Z、?、;、,、.、!、)、(、%……à、â、è、ê、é……然后，我们就可以把所有以相同字符开头的书摆在同一楼层。

现在我们的图书馆只有100层高，比芝加哥的西尔斯大厦还低一些。我们可以把每一层楼划分成100个走廊，并按字母表的顺序标记这些走廊，每个走廊中书的第二个字符都相同。接着，我们在每个走廊放置100个书架，每个书架用来存放第三个字符相同的书。因此，所有开头为"土豚喜欢莫扎特"（aardvarks love Mozart）的书都将摆置在第一层楼第一个走廊的同一个书架（"r"架）上。但这会是一个非常非常长的书架，也许，我们最好把书放在与书架成直角的抽屉中，每个抽屉用来存放第四个字符相同的书。这样的话，每个书架大概就只有30米长了。但是，现在每个放书的抽屉都深不见底，而且还会抵到相邻走廊抽屉的背面，所以……但我们已经用尽了所有的维度去安排这些书。

我们需要一个100万维的空间将所有书整整齐齐地存放起来，但实际上我们只有三个维度：上下、左右和前后。因此，就不得不假装我们

可以想象一个高维空间，其中的每个维度都与其他维度“垂直”。即使我们不能形象化地表现它们，也还是可以设想这样的超空间（hyperspace）。一直以来，科学家们利用它们为自己的理论表达赋予生机。无论它们是否只是想象，这种空间的几何结构表现优异，数学家们对此也进行了深入的探索。所以，在这些逻辑空间中，我们就可以安心地谈论位置、路径、轨迹、体积（超体积）、距离以及方向。

现在，我们接着考虑博尔赫斯主题的一个变种，我打算称之为孟德尔图书馆。这个图书馆包含“所有可能的基因组”——DNA序列。理查德·道金斯在《盲眼钟表匠》（*The Blind Watchmaker*, 1986）中描述了一个类似的空间，他称之为生物形态园（Biomorph Land）。他的讨论激发了我的灵感，我们俩的解释完全相容，不过我想强调一些他轻描淡写的内容。

设想孟德尔图书馆是由对基因组的描述组成的，那么该图书馆就恰好是巴别图书馆的一部分。描述DNA的标准代码只包含四个字符，A、C、G和T，分别代表腺嘌呤、胞嘧啶、鸟嘌呤和胸腺嘧啶四种核苷酸。因此，所有用这四个字母填满500页的排列已经在巴别图书馆里了。但是，典型的基因组要比普通书籍长得多，人类基因组中大约有30亿个核苷酸，因此一个人的基因组，比方说你的基因组，就将占用巴别图书馆里的大约3 000本书。

现在，人类的基因组与《白鲸》星系的这种比较对人类基因组之间的差异性和相似性做出了解释。如果人类个体的基因组各不相同，不只是在某一处不同，而是在成百上千的位置（用遗传学的话说，叫作基因位点）上都有所不同，我们又何谈为全体人类基因组测序呢？就像大家熟悉的雪花、指纹那样，没有两个人类基因组是完全相同的，包括同卵双胞胎的基因组，遗传代码变异的机会总是存在的，即使对于同一个个体的细胞来说也是如此。人类DNA可以毫无困难地同其他任何物种的DNA区别开来，即便对于黑猩猩也同样如此，虽然它与人类有90%以上的相同位点。

所有存在过的某个人类的基因组都被包含在一个可能的人类基因组的星系中，它同其他物种的基因组之间遥不可及，然而，该星系有足够的空间使任何两个人的基因组都不一样。你的每一处基因都有两个变体，一个来自你的母亲，另一个来自你的父亲。他们恰好将自己一半的基因传给你，而他们从他们的父母，即你的祖父母和外祖父母那里通过随机选择获得了这些基因，但因为你的祖父母和外祖父母也是智人中的

成员，他们的基因组几乎在所有位点上都一样，因此，在绝大多数时候，不管你的哪一位先祖为你提供了基因，都没有差别。但他们的基因组仍然在成千上万的位点上不同，而在这些有差异的位置上，你得到哪个基因完全是偶然的，就像在双亲对你DNA的贡献过程中内置了一个掷硬币机制。此外，在哺乳动物中，突变积累的速率大约是每代每个基因组100个突变。“也就是说，你的孩子与你们夫妻二人的基因会有100处不同，这是酶的随机复制错误产生或宇宙射线造成卵巢及睾丸突变的结果。”（Ridley, 1993, p. 45）

对马、兔子或章鱼的基因组描述也由同样的字母，即A、C、G和T组成。大多数动物的基因组比人类基因组小，但是一些植物的基因组要比人类的基因组大10倍以上，而一些单细胞生物的基因组还要更大！单细胞生物变形虫（amoeba dubia）是当前世界记录的保持者，据估计，它的基因组拥有6 700亿个碱基对，是人类碱基对数量的200倍还要多。但让我们姑且假定孟德尔图书馆都是由基因串的描述构成，而且，这些描述写在一套全部由四个字符印制而成的3 000册丛书上。这样这个图书馆将包含足够多“可能”的基因组，完全能够满足任何用意严肃的理论需要。

在描述孟德尔图书馆时，我夸张地说它包含“所有可能的基因组”。然而，正如巴别图书馆忽视了俄罗斯语和汉语等语言，孟德尔图书馆也忽视了显然还存在其他遗传系统的可能性，例如，基于其他化学构造的遗传系统。所以，一旦我们试图将得到的任何结论应用在一些更为宽泛的可能性概念之上时，我们也许就不得不重新考虑对这座孟德尔图书馆来说什么是可能的。这实际上是一项策略上的优势，而非缺陷，因为我们可以对我们所谈论的那种适度的、有所限制的可能性保持密切关注。

DNA的重要特征是，腺嘌呤、胞嘧啶、鸟嘌呤和胸腺嘧啶的所有排列都具有同样稳定的化学性质。原则上，在基因拼接实验室，所有碱基组合都有可能被构造出来，而且一旦将其组装出来，它们将具有无限保质期，就像图书馆里的书一样。但在孟德尔图书馆，并非每一个这样的序列都对应着一个可行的生命体。大多数DNA序列都是乱码，毫无生机可言。所有我们见到过的基因组，那些到今天还实际存在的基因组，是几十亿年以来调节和修正的结果，是盲目编辑过程的产品，这一过程之所以有效，是因为大部分的乱码会被自动放弃，而只有数量微渺，但有意义、有用的“文本”会留存下来，翻来覆去地被使用、被复制。

就在现在，你身上就有超过一万亿份你的基因组副本，每个细胞中

均有一份，而且在每一天，随着新的皮肤细胞、骨细胞和血细胞被制造出来，新的基因组副本也被安置于其中。可被复制的文本得以复制，因为它驻留在持续运转的活细胞中，而其余的则被消解掉——要么发表，要么灭亡（publish or perish）。

36

什么是基因

基因就像软件中的子程序

正如我们刚才看到的，将基因类比为字词非常有用，但还有一个更好的类比，我们借助这段有关计算机的“插曲”就能明白。理查德·道金斯在其杰作《祖先的故事》（*The Ancestor's Tale*, 2004）中就进化论问题对另一位才华横溢的作家马特·雷德利（Matt Ridley）大加赞赏，缘由是后者在他的书《先天后天》（*Nature via Nurture*）中指出基因和软件子程序之间深刻的相似性。我通常不会在自己的书中大段大段地援引他人的文字，不过我发现，如果只是为了一丁点原创性，就去改述这段话的话，那么我总是会牺牲掉明澈性与生动性。所以，经道金斯许可，我沿用了他的原始表述。

> 我们所测定的大多数基因组并不是一套用以创造一个人或一只老鼠的指令集，也不是什么计算机主程序，尽管有一部分基因组充当这种功能。如果真是这样的话，我们确实会预期我们的程序比老鼠的大。但大多数的基因组更像是可用于写出这套指令集的词典，或者说像是被主程序调用的子程序。如雷德利所言，《大卫·科波菲尔》和《麦田里的守望者》中用到的单词几乎完全一样，两者都出自一个受过教育、以英语为母语的人的词汇表。这两本书的不同之处在于这些单词连接在一起的顺序。
>
> 当一个人或一只老鼠形成的时候，两者的胚胎均借鉴了一部相同的基因词典：哺乳动物胚胎学标准词汇表。人和老鼠的区别在于从共有的哺乳动物词汇表中所引用的基因的不同顺序，以及这些基因在身体的不同位置与时间表达上的差异。这一切都在某些特殊基因的掌控之中，这种基因的作用是以一种复杂精妙的定时级联的方式发动其他基因。但是，在整个基因

组的构成中，这种主控基因只是很小的一部分。

> 不要将这里所说的顺序误解为基因在染色体上排列的顺序。除了极个别的例外以外，……这些基因在染色体上排列的顺序就像单词在词典里排列的顺序一样，通常是按照字母顺序排列的，但有时也会按照方便使用的顺序排列，就像给国外游客用的小册子，按照机场常用词汇、就医常用词汇、购物常用词汇等排列。基因在染色体中存放的顺序并不重要。重要的是细胞在需要的时候启动正确的基因，而这其中的机制我们正理解得越来越多。
>
> 在某一方面，这种基因和单词的类比是具有误导性的。单词比基因更短，所以有些作家会将基因比作句子。但出于其他原因，句子也不是一个很好的类比。不同的书并不是由固定不变的句子排列而成的，大多数句子都是独一无二的。基因更像字词而不像句子，它可以在不同的上下文中反复使用。比单词或句子更好的类比是将基因比作计算机中的工具箱子程序……
>
> 苹果电脑的只读存储器或启动时加载的系统文件中存放着程序工具箱。这些成千上万的工具箱子程序各自执行特定的任务，它们在需要时以不同的方式在不同的程序中被反复调用。例如，模糊指针（ObscureCursor）这一工具箱子程序的作用是在鼠标下次移动之前将屏幕上的指针隐藏起来。每次当你打字的时候，系统就会调用模糊指针这个“基因”，将鼠标指针藏起来。工具箱子程序支持着苹果电脑以及在Windows系统中与之对应的东西中所有程序共有的特征：下拉菜单、滚动条、在屏幕上拖动鼠标改变窗口的大小等。（Dawkins, 2004, pp. 155-156）

所有这一切都有助于让我们了解到专家为什么能如此轻松地识别出哺乳动物的基因组。原因是，我们拥有哺乳动物的工具箱，它里面除了专门制作哺乳动物的工具外，还包含来自爬行动物、鱼类，甚至蠕虫工具箱中的工具。这套工具中最古老的工具为所有生物所共享，其中甚至包括细菌。

37

生命之树

我们人类与谁有着共同的祖先

现存基因组由世系谱线连至其父系、祖系等的基因组，可以追溯到地球上的生命之初。本书插页中的生命之树，显示出每个人同其他人的关系有多么密切：人类在10万年前有共同祖先，在2亿年前与狗和鲸有共同祖先，在20亿年前与菊科植物和红杉有共同祖先。

绘制生命之树有多种方式。在目前这种绘制方案中，各个分支外缘的顶端代表现在，也就是说，只有到达外缘的世系存活到了现在。该图表明，恐龙早在6 000万年前就已灭绝，而它们的后裔鸟类则存活至今。所有世系的谱线最终都连回到生命之初，同时也连至后来者。如果将该图放大一万亿倍，我们就可以看到曾经存活过的每一只两翼昆虫、每一条鱼以及每一只蛙的系谱，看到它们当中的大部分无后而终，有的则繁衍至今。

38

自然选择VS智能设计
“起重机”VS“天钩”

生命令人称奇。想想这宇宙间有无数几乎完全死气沉沉的恒星系统，我们便会惊讶竟然还有生命的存在。同样令人称奇的是那些形形色色的生存之道，从细菌到鱼类、鸟类、雏菊、鲸、蛇、槭树，再到人类。也许最为神奇的是生物的坚韧不拔：它们千方百计地顽强生存、繁衍，面对巨大的阻碍竭力谋生。这一切都要归功于数以百万计的巧妙装置和安排：从细胞中蛋白质机器错综复杂的级联，到蝙蝠的回声定位、大象的鼻子，再到我们的大脑所具有的能够反思“天底下”所有问题的能力等。

达成这一切的鬼斧神工需要解释，因为这不可能是出于纯粹的机缘巧合或偶然。已知只有两种可能性：智能设计或经由自然选择的进化。无论是借由智能设计者奇迹般的设计，还是通过自然选择步履笨拙、毫无先见、愚蠢且非奇迹的进化，两种方案都需要极大量的设计工作。我们称其为研发（resaech and development）。研发的代价总是巨大的，且需要时间与精力。达尔文伟大想法的美妙之处在于，他看到了历经几十亿年的运作和数量惊人的“无用”功（不计其数的尝试以失败告终），设计改进如何以一种自然而然、非奇迹的方式累积，不需要任何目的、预见和理解。达尔文最热心的批评者罗伯特·贝弗利·麦肯齐（Robert Beverley MacKenzie）曾雄辩地抨击道：

> 对于这种理论，我们必须得有一个绝对无知（Absolute Ignorance）的设计者，以便我们可以将其解释为整个体系的基本原则，即：要想创造出一架完美出色的机器，我们并不需要知道如何去造。细加检讨，不难发现，该命题凝练地表达了进化论的宗旨，只言片语就说出了达尔文先生的所有想法：经由一种奇怪的推理倒置，他似乎认为，在创造性技能的所有成果中，绝对无知完全有资格取代绝对智慧（Absolute Wisdom）。（MacKenzie, 1868）

没错，自达尔文提出这个惊人的想法之后，怀疑者们就质疑是否有足够的时间来完成所有这些创造性的工作。如何设想所需完成的设计工作呢？一个简单的决窍是将其想作设计空间中的提升。什么是设计空间？最好将它想象为一个多维空间，就像巴别图书馆或孟德尔图书馆那样。事实上，设计空间不但包括这类图书馆，而且还包括其他图书馆，因为它不仅包括所有经由构思而创作的书籍和经由设计而进化的生物，而且还包括由设计立场所描述的其他物事（见第18章），例如，房屋、捕鼠器、战斧导弹、计算机和宇宙飞船。正如巴别图书馆中的大多数书籍是乱码，设计空间中的大多数处所也都充满了毫无用处的垃圾。如果你和我一样，只能一次想象三个维度，请相信我，你越是在想象中摆弄这种想法，就越容易以这种熟知的三维为基准想象出更多的维度。这是一种思考工具，你用得越多就越有提高。

设想把生命之树置于孟德尔图书馆，我们便可以看到，所有曾经存在于这个星球上的生灵，它们的世系如何彼此联系。这些世系传递着自然选择所“发现”的基本的设计改进，为所有后继生命保留了最初的研发成果。从细菌到脑细胞，每一个细胞都拥有这种“机械装置”，包括不计其数的、已经运转了超过30亿年的精巧的纳米装置，那是我们同树木、鸟类和酵母细胞所共享的生命的基本引擎舱。在多细胞层面，我们可以看到心脏、肺、眼睛、手臂、腿以及翅膀，它们“仅仅”在20亿年前才被首次研发出来，并经过了重新利用和改进。但是，除了生物的器官，还有它们的造物，例如，蜘蛛网、鸟巢以及河狸修建的水坝。这些东西同样也是研发确凿无疑的证据，而且研发的丰富成果也以某种方式一脉相承，不单在基因中传递，同样也借由子代复制父代这一过程传递。新奇事物通过突变、试验或意外不时地涌现出来，这是一种改进，它一次又一次地得到了复制。而失败的试验品走向了灭绝。再说一次，要么发表，要么灭亡。

再有就是人工器物：耕犁、桥梁、教堂、绘画、戏剧、诗歌、印刷机、飞机、计算机、剪草机等，以及思考工具。它们是怎么回事？难道它们不属于生命之树中人类世系这一分支吗？它们不仅仅依赖于一个作者或发明者，同样也极大地依赖于背后众多的研发贡献者们。贝多芬不必发明交响乐，因为它已经存在了，同样，莎士比亚也没有必要创制十四行诗。一个链锯由几十或几百个“现成”的元素构成，它们已经被发明出来了，而且已经得到了优化。人类的一些器物很可能复制了其他动物所创造的东西。织巢鸟的巢有没有激发出织布工艺的发明呢？因纽特人圆顶冰屋里的活动地板可以使冷空气从地下出口排出，是不是模仿了北

极熊窝里类似的活动地板呢？还是说这些是独立自主的发明？或者是北极熊模仿了因纽特人？

心脏与水泵、蝙蝠的回声定位与声纳和雷达、河狸坝与灌溉水坝、眼睛与相机，它们之间在功能方面深层次的相似性并不出人意料。在过去的许多个世纪中，同样的研发探索过程已经将它们打磨成形，并不断改进。它们可以被放置在同一个设计空间里，即所有可能的设计空间。

这一说法在生物学家中间是有争议的，我自认为理解个中缘由，并对此深表遗憾。许多生物学家都极不愿意谈论对生物的“设计”，因为他们认为这会为智能设计运动提供某种支援或慰藉，而智能设计论是不真诚的，是伪科学，是某种隐秘的“宗教”，有损进化生物学当之无愧的权威。在这些抵制人士当中包括两个我最为敬重的同事兼朋友：进化生物学家理查德·道金斯和杰里·科因（Jerry Coyne）。通过讨论自然的目的和设计，我们显然为支持智能设计的那帮人提供了半数案例。有些人认为最好对此类主题严厉封杀，他们坚持认为，严格说来，生物界不存在任何被设计的东西，除非它是由人设计的。大自然用以生成复杂系统，例如器官、行为等的方式与设计者所采用的方式非常不同，我们不应该使用相同的语言描述它们。所以，道金斯偶尔才会谈到生物类设计的特征。在《祖先的故事》中，他说道：“由达尔文的自然选择学说带来的设计幻象具有一种鼓动人心的力量。”我对这种紧缩策略并不认同，因为它很可能适得其反。

近日，我无意间听到酒吧里的一群年轻人在谈论细胞中发现的纳米机器。其中一位感慨道：“看看这些奇妙的小机器人，你怎么能相信进化论！”而其他人若有所思地对此点头称是。不知怎的，这些人得到的印象竟然是进化生物学家认为生命并不复杂，构成生命的组件也不是奇妙无比的。这些进化论的怀疑者可不是不学无术的人，他们是哈佛医学院的高材生！他们极大地低估了自然选择的力量，因为进化生物学家三番五次地告诉过他们，自然界不存在任何真正的设计，它们只是看起来像设计。这个故事强烈地提示我，“常识”中已经开始接纳了一种错误的想法，即进化生物学家不愿意“容许”或“承认”自然界中存在显而易见的设计。

说到这里，让我们来看看克里斯托夫·舍恩博恩（Christoph Schönborn）的话，他是天主教的维也纳大主教，被智能设计那帮人愚弄了的家伙。在《纽约时报》的专栏文章《寻找自然中的设计》（*Finding Design in Nature*）中，他写了一段众所周知的话。

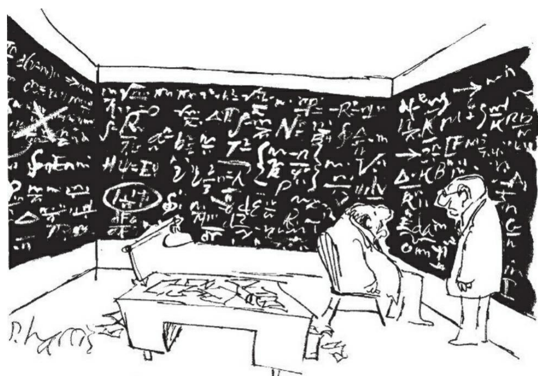
> 在把地球生命史的诸多细节托付给科学之后，天主教教会依然可以宣称，凭借理性之光，人的理智可以很容易、很清晰地领悟包括生物世界在内的自然世界的目的和设计。在共同祖先的意义上，进化可能是真的，但新达尔文主义认为进化是无导向、无计划的随机变异和自然选择过程，这并不正确。任何否认或寻求压倒性的证据来消除存在于生物之中的设计的思想体系都是意识形态，而非科学。

我们这些进化论者想要将论战带向何方呢？是要试图说服门外汉，让他们相信他们看到的那些存在于每一个生物尺度上的显而易见的设计并不真实，还是该努力表明，达尔文只是揭示了真正的设计也可以无需借助智能设计者之手？我们已经说服全世界相信地球绕着太阳转，时间是相对的，不是绝对的。可为什么当我们试图表明可以存在没有设计者的设计时，就变得畏缩不前了呢？所以，在这里我要再次捍卫以下主张。

生物界充满了设计、目的和理由。我所说的设计立场预见并解释了贯穿生物世界的全部特征，同样的假设在对多少有点聪明的人类设计师所做的设计进行逆向工程时运作良好。经由自然选择的进化是一系列过程，该过程“发现”并“追踪”了事物被如此安排而非另有安排的理由。进化过程所揭示的理由和人类设计者所发现的理由之间的主要区别是，后者通常（但并不总是）会在设计者的头脑之中被表达出来，而前者一般在人类探索者首次成功地对自然造物进行逆向工程之后才会被表达出来。也就是说，人类设计者会思考其造物之所以具有某种特征的理由，并因此有了表达这些理由的想法。他们通常会欣赏、制定、细化、传达、讨论、批评这些设计理由。进化过程并没有这么做，只是通过它产生的变异盲目地进行筛选，使好的东西得以复制。这些东西之所以是好的是有理由的，但自然选择过程并未表达这些理由。

用类设计（designoid）去标识那些仅仅看上去像是设计的东西可能颇有助益，但在生物学上并非如此。当我一想到那些仅仅看上去像是设计的东西，我立刻就会联想起漫画家设置的一些场景：大胡子理论家们或站在布满了实际上毫无意义的符号的黑板前，或置身于一间化学实验室，其中堆满了令人印象深刻的形形色色的试管和烧杯，还有疯狂发明家在摆弄时间机器，上面天线林立，密布各式各样的仪表盘和高科技玩意儿。这些东西不能做任何实际的工作，它们只是看上去有用。但是，大自然的设计是真正有效的。事实上，它们通常要比任何人工设计更为

有效和强大，能与之匹敌的人工设计还没发展出来呢。



就没有一个优雅的解决方法吗？

一个身体健康的年轻人携带所需的食物和水，一周大约能步行240公里。水是关键负重：在背包中携带23千克的水、7千克的食物以及5千克的装备会使之感觉非常沉重。若能在路途中找到水，这个人也许能坚持步行好几个月。作为比较，让我们看看康奈尔大学的机器人专家安迪·鲁伊纳（Andy Ruina）和他同事的心血：迄今为止行进距离最长的步行机器人"游侠"（Ranger），这部机器人曾在日本2001年5月1日至2日举办的机器人超级马拉松大赛上马不停蹄地走了65.2公里。游侠的设计者们利用肢体动力学创造了一个强健的高能效步行机器人，可以被摇杆遥控着在平坦的赛道上一圈又一圈地行走。另外一个出色的四足步行机器人是"大狗"（Big Dog），它的能源利用率比游侠低15倍，但更令人印象深刻的是它适应复杂地形的能力。尽管如此，人体传动装置的效率仍然是游侠动力效率的四五倍，还有，人类不像游侠，他们可以自主地对这个世界上形形色色的地理特征做出回应（Ruina, 2011）。

看来，我们必须放弃某种东西。要么将"设计"定义为某个智能设计者的作品，例如诗歌、汽车；要么承认有可能存在无需假借智能设计者之手的真正设计。传统以及词源上的解释似乎偏向于前者，但考虑一下这种情况："atom"一词来自希腊语a加上tomos。a意为"无"；tomos来自temnein，意为"切割"或"切开"。原子一词最初的意思是"无法切开的东西"，然而后来科学发现原子可分，这与其定义并不矛盾。我认为，科学同样也揭示出没有一个有头脑、有先见之明、有意向的设计者的设计不但是可能的，而且就在我们身边。经由进化做出的设计是一个真实的、易于理解的过程。它在一些有趣的方面不同于工程师的设计，但它们也

高度相似，都具有震撼人心的“独创性”。一次又一次，生物学家对自然界中那些看似无用或有点笨拙的糟糕设计颇感困惑，但最终意识到自己低估了自然造化的神妙。弗朗西斯·克里克调皮地将这种趋势冠以他同事莱斯利·奥格尔（Leslie Orgel）的名字，即所谓的奥格尔第二定律：“进化比你们聪明。”将自然选择过程拟人化地比作大自然母亲这一策略显然是考虑不周的，但我们将会在适当的时候为这种挑衅性的人格化说法进行辩护。它不只是一个有趣的隐喻，其本身就是一种思考工具。

下面，让我们回到设计空间，这个包含所有可能设计的多维空间不仅包括实际的生物、汽车和诗歌，而且和巴别图书馆一样，包含着所有从未进化、从未被构造出来的设计，比如，能说话的袋鼠、会飞的蛇、核动力爆米花桶和水下轮滑鞋。基础设计元素的词汇表应该是什么样的呢？

我们并不打算造一个设计空间出来，只是要设想一下，所以我们尽可以肆无忌惮。比方说，设想词汇表是那些来自元素周期表中原子的所有可能组合。这个浩瀚空间中甚至包括所有并非设计出来的东西的复制品，比如，海滩上的每一块鹅卵石以及珠穆朗玛峰，因为没有任何东西阻止某人去设计或建造这些实体的复制品。那么，勃拉姆斯第三交响曲会在设计空间的什么地方呢？乐谱以纸上的墨迹等形式会出现在很多地方，磁带和唱片又会出现在其他地方，因此，它一定在设计空间之中。只通过口述留存和传诵、从未被记录下来的歌曲很难被归置于上述原子家族，但是，随着时间的流逝，一个足够复杂的设计空间的子空间将可以容纳它们。

正如巴别图书馆的绝大部分地方摆满了毫无意义的书籍，设计空间里也充满了无意义的东西，这些东西毫无生趣、缺乏用途，也没有任何作用，但不经意间，真实或可能设计的那微渺的线索便会闪烁一缕微光。万物不是仅仅在热力学第二定律的无情法令下静候灭亡的，它们自有其作为。

借助设计空间这一粗糙的想法，现在我们就可以“一窥”自从达尔文时代就一直困扰进化论的那些争论是如何成形的：无论是自然的还是人工的，是否存在一种真正的设计，它既非直接、又非间接地来自那棵单一的生命之树？

以下是一些备选答案。

（1）不存在。

（2）存在。一些自然的奇迹太奇妙了，有着“不可还原的复杂性”，而单调乏味的进化设计过程根本不可能造就这样的奇迹。它们必定是由一个智能设计者创造出来的。

（3）存在。一些人造的东西，比如说，莎士比亚的戏剧、哥德尔定理，它们太过神奇，不可能“仅仅”是经由进化而来的人类大脑的产物。它们是不可思议的天才之作，单调乏味的进化设计过程无法理解它们，绝对无法企及。

我们可以认为（2）（3）这两个回答高度依赖于这一对比：完全低效、单调乏味的自然选择过程（麦肯齐所谓的“绝对无知”）和莫扎特轻快灵动、不费吹灰之力的卓越才智（或者，你也可以换成任何一位“神一样”的天才）。第49章将探讨这个问题。借助我们在提升和研发之间做出的类比，可以看到，这两个答案都需要天钩（Skyhook）。《牛津英语词典》对“天钩”的解释如下：

> 天钩：源自飞行器驾驶员。一种想象中联结至天空的装置；一种想象中悬停在空中的方式。

《牛津英语词典》记录该词最早出现于1915年：“一名飞行员被要求停留在原位置（高空处）一小时，飞行员答道：‘飞机未配备天钩。’”天钩概念也许是古希腊戏剧学中天降救星（deus ex machina）[(71)] 的衍生物：一旦一个二流剧作家发现剧情让他们的主角进退维谷，通常就乐于在场景中设计降下一尊神以力挽狂澜，就好比超人那样。或者，天钩可能完全是诸多民间传说独立演变而来的产物。拥有天钩可能会是件美事，因为它非常适合在困难的情况下起吊笨重的物体，用以加快各类项目建设。不过，说来让人失望，它们是不可能的。[(72)]

还好我们有起重机。起重机可以做想象中天钩所做的提升工作，而且它们可以以实际的、非循环论证的方式完成这种工作。它们很昂贵。我们必须从手头上已有的寻常部件开始设计和建造它们，它们也必须坐落于一个真实存在的坚实基础之上。天钩是件神奇的升降装置，无根无据。而起重机是件毫不逊色的升降装置，且具有真实可行的优势。任何像我这样一直在建筑工地旁边观察的人都会相当满意地注意到，有时候我们需要用一部小型起重机装配一部大型起重机。许多其他旁观者也必定注意到，原则上，这架大型起重机可用于启用或加快更大型的起重机的安装。起重机的级联是一种策略，在现实世界的建设工程中我们很少

会采用这种策略，但原则上，用以级联的起重机的数量是没有限制的，我们可以用它们来完成一些宏伟的工程。[73]

现在想象一下，要想创造出我们在这个世界上看到的诸多壮丽的生物体以及其他造物，必须在设计空间中完成所有的“提升”。自生命曙光乍现，从最早期、最简单的自我复制的实体开始，造物必须穿越巨大的距离，向外（多样性）、向上（优越性）扩散。达尔文为我们提供了一种最原始、最基本，同样也是最笨拙的提升过程：自然选择的斜面。经由极为微小的步骤，这一过程在极长的时间内能够逐渐跨越这些巨大的距离。至少达尔文是这样说的。在任何地方都不需要天降奇迹。每一步都始于早期攀爬的诸多努力所营造的基础，经由粗糙、机械、系统渐进式的攀爬而实现。

这一切看似令人难以置信。它真的能发生吗？或者，这个过程会时不时地需要某种天钩“帮一把”？抑或只在非常早的时期被帮了一把？一个多世纪以来，怀疑者一直在试图证明达尔文的观念并非有效，至少不是自始至终有效。他们一直希望觅得天钩，这是他们眼中前景黯淡的达尔文渐进系统的例外。他们屡屡提出了一些确实很有意思的挑战：跳跃、断层和其他奇迹，乍一看这些似乎都需要天钩。但是，起重机随之而来，在许多场合，它们正是被那些希望找到天钩的怀疑者们发现的。

是时候做一些更细致的定义了。这么解释吧，天钩是一种“心智至上”的强大力量或过程，是“所有设计以及看似设计的东西最终都是无心智、无动机机械过程的结果”这一原则的例外。相反，起重机是一个设计过程的子过程或其特殊性质，被证明可以对自然选择这一基本、缓慢的过程进行局部加速，而且起重机本身就是这一基本过程可预测的或回过头来能被解释的产物。一些起重机是显而易见、没有争议的，其他一些起重机则还处于激烈的争论中。为了让我们对这个概念的广度和应用有个一般的概念，让我举三个非常不同的例子。

“共生”是部起重机。我最喜欢的例子之一就是真核生物的共生起源。如果看看生命之树，你就会注意到，多细胞生物波澜壮阔地呈扇状展开，其中包括所有动植物，它们的出现时间约在真核生物之后。大约十亿年间，这个星球上唯一的生物只是单细胞生物：细菌和古生菌，统称为原核生物。然后，某个幸运的一天，一个原核生物碰巧撞上了另一个原核生物。毫无疑问，这一过程发生得非常频繁，时至今日仍然如此。但那一次，产生的结果并不是其中的一个生物吞噬或分解掉另一个，也不是相互排斥，而是它们强强联手了，用大致两倍于各自的数量

和更多种类的工作部件制成了另一种全新的、有生命的东西。两个研发世系曾缓慢地各自进化了数百万年，磨炼了其独有的能力，现在它们走到了一起，并最终一“战”成名。“技术转让”并不总能产生良好的效果，这种事不常发生，可是一旦发生，结果便会蔚为壮观。两个世系中的任何一个都犯不着重塑另外一个世系的生存策略和体系，因为它们各自能力的结合正好是净收益，流行的说法叫“协同增效”，真核生物比它们各自任何一种都更适于生存，所以真核生物的世系繁盛起来。

以前，人们对于真核生物的起源是一个共生事件的主张是有争议的，但现在有非常多的证据支持这一主张，而且该主张已经稳稳地写入了教科书。若能证明细菌鞭毛的起源得益于共生这部起重机，那就令人心满意足了，毕竟这是智能设计论支持者长期坚持的“不可还原的复杂性”的招牌例子。然而，已故生物学家、真核生物共生起源的拥护者林恩·马古利斯（Lynn Margulis）曾不遗余力地捍卫这一主张目前的证据支持对鞭毛的起源做出另一种进化解释。鉴于这么多丰富的证据，声称鞭毛的起源需要天钩显得特别凄凉。

看一眼生命之树便知，真核生物为所有多细胞生物的出现做好了铺垫。细菌和古生菌右侧那些多姿多彩的世系都属于真核生物。大致而言，能被肉眼看到的生物都是真核生物。真核革命为设计空间开辟了巨大的领域，但它的发生并不是为了让所有这些设计能够出现。起重机必须在局部就能带来好处，它所传递的那些设计革新要带来直接收益。但是一旦这些起重机建立起来，就会有更深远的影响。同样的道理，电脑的发明并不是为了使用文字处理器或互联网，但是，一旦进入了一个可能的计算机应用程序空间，设计过程便开始超速运转，创造了我们现在每天所依赖的所有“物种”。

性也是一部起重机。现在，进化理论家们普遍认为性是一部起重机。有性繁殖的物种可以比无性繁殖的物种更快地穿越设计空间。此外，它们还可以“识别出”进化进程中的设计改进，而无性繁殖的物种对此“毫不知情”（Holland，1975）。然而，这不可能是性存在的理由。进化过程是盲目的，所以，任何它所建立的东西必须有立竿见影的回报以抵消成本。

正如最近理论家们强调的那样，“选择”有性繁殖需要承担巨大的直接成本：在任何一次交易中，生物体只能将它们基因的50%传给下一代，更不要说确保交易成功需要付出的努力以及交易进行所面临的风险。因此，效率、敏感度和再设计速率的提高虽然使性成为一部宏伟的

起重机，但这些特性所带来的长期回报对短视的局部竞争来说一无是处，而正是局部竞争确定了下一代当中的哪些生物体更能获得青睐。必然有其他一些短期利益维持着有性生殖所需的正向选择压力，使得很少有物种可以拒绝这一选择。生物学家约翰·梅纳德·史密斯（John Maynard Smith）最早强有力地提出了多种多样的引人注目且相互竞争的假设，这些假设可能会解决这一难题。对于进化生物学中这一重要议题的清晰介绍，可以看看马特·雷德利的《红色皇后：性与人性的进化》（*The Red Queen: Sex and the Evolution of Human Nature*, 1993）。

基因工程也是一部起重机。性揭示了拥有超凡威力的起重机的存在也许并不是为了彰显这一力量，尽管作为一部起重机，它的力量可能有助于解释它为什么会一直延存至今，但它的存在却是出于其他理由。显而易见，基因工程是被当作一部起重机而创造出来的。毫无疑问，现在基因工程师可以穿越设计空间实现巨大的飞跃，创造出用“平常”方式绝不可能进化出的生命体。基因工程师以及他们在操作中使用的各种器物本身就是早前缓慢进化过程的产物，因此这并不是奇迹。

但如果神创论者是对的，即人类这个物种本身是神圣的，无法经由简单粗暴的达尔文式路径达到，那么基因工程将根本不是一部起重机，要创造出这部起重机还需借助某个至关重要的天钩。我无法想象有基因工程师会认为自己是用这种方式被创造出来的，但该想法确实是一个有效的逻辑支点，虽然这个支点并不可靠。还有另一种说法看上去不那么愚蠢：如果基因工程师的身体是进化的产物，但其心智能够完成创造性的工作，而这些工作是不能再简化的非算法形式，也无法通过任何算法路径到达，那么基因工程的飞跃就很可能涉及天钩。稍后我们将简要地探讨这一图景。

39

小布谷鸟为什么会把宿主的蛋推出鸟窝
无需理解的能力之一

绝对无知是“创造性技能所有成就”的源头。达尔文的这一想法被麦肯齐描述为“奇怪的推理倒置”，因为它颠倒了一个再“明显”不过的想法：理解是能力的来源。我们为什么要送孩子上学？为什么要强调“理解概念”而不是“死记硬背”？因为我们认为，在任何领域的活动中，理解是能力的最佳达成路径。不要满足于做一个没头脑的人！无论我们在做什么，只要理解了事情的原则，就可以做得更好！对于人类活动的大部分领域来说，这无疑是条很好的忠告。但我们也发现存在一些极端的特例，比如，某些天才音乐家不识乐谱，但听到一段音乐就能演奏出来；某些天生的运动员似乎总是表现出色，却不能解释他们为什么这样做、是怎么做到的，因此也不能指导别人。还有所谓的“白痴专家”，他们在很多方面的能力都很欠缺，而在某些有限的领域中却能力超凡。但总的说来，经验法则无法否认：理解通常是人类能力的关键。

达尔文确实倒置了整个推理，正如麦肯齐颇为生动地指出的，他表明绝对无知是一个能工巧匠。自然选择过程具有惊人的能力（想想奥格尔第二定律），但却是全然无知的。此外，生物从它们的设计带来的精良装备中获得了诸多好处，而无需理解它们为什么有此天赋、天赋是怎样运行的。我最喜欢的例子是布谷鸟。布谷鸟是种巢寄生鸟类，它们自己不筑巢。相反，雌性布谷鸟偷偷将蛋下在其他鸟类的巢中，在那里，布谷鸟的蛋等待得到被蒙在鼓里的养父母的关注。通常，雌性布谷鸟会将一个寄主的蛋从鸟巢中推出去，以防寄主发觉！当小布谷鸟孵出来时，由于它往往要比寄主的蛋孵化得快，羽翼未丰的它们会努力将剩余的蛋推出鸟巢！为什么？为了最大限度地获得养父母的关注。

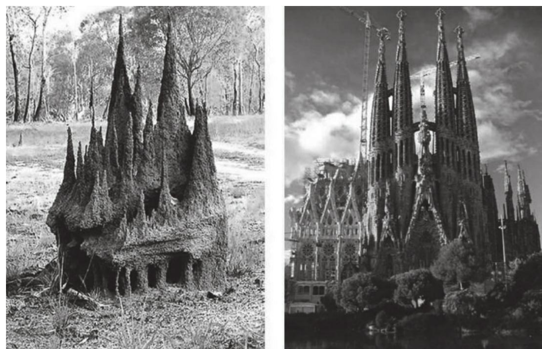
40

白蚁城堡是谁设计的
无需理解的能力之二

自然选择是一个自主的理由搜寻者，它历经多个世代“发现”“认可”，并“集中”了诸多理由。这些引号是为了提醒我们，自然选择并没有心智，其自身也没有理由，但它却可以胜任设计改进的“任务”。这本身便是一个无需理解就能胜任的实例。我们要知道如何落实这些引号。考虑一下，某个种群内部有大量的变种。其中一些成员活得不错，也善于繁殖，而大多数成员并不精于此道。对于每一种不同的情况，我们都可以问为什么。为什么这个个体有后代，而其他个体却没有？在大多数情况下，根本没有任何理由可言。它只是撞大运而已。但是，如果该种群中有一个子集，也许是一个非常非常小的子集，在这个子集中，某种个体间的差异会有所作为，那么，这些共同的差异就会萌发出某个理由。这就使得功能性得以通过一个盲目跟进理由的过程得到累积，从而创造出带有目的的东西，而且不需要知道目的的存在。需知原则（The Need to Know）主宰了整个生物圈，而自然选择过程本身并不需要知道它在做什么。

因此，在有能够表述理由的个体之前就存在着理由了。我将这种演变过程所跟进的理由称为“自由浮动的理据”（free-floating rationales），这一术语显然刺痛了某些思想家的神经，他们怀疑我变出了某种幽灵、某种奇怪的非物质理念，它没有理由出现在一个冷峻严肃的唯物主义者关于实在的解释之中。实情当然不是这样。自由浮动的理据和重力、重心一样，并不是幽灵或有问题的。在人们发明出清晰准确的算术方法前，就已经存在着九大行星；[74]在物理学家想到重心的概念和计算方法前，小行星就有重心。将数等同于阿拉伯数字、罗马数字或其他什么数字是一个错误，数字不过是我们所使用的数的名字。数字是人类的发明，而数不是。我所使用的“理由”这一术语更像数，而不是数字。在人类探索者或其他心智表述、体现这些理由之前，它们已经由进

化过程所揭示。让我们仔细看一看下图中两个引人注目的相似结构。



蚁冢和圣家族大教堂

蚁冢和高迪设计的圣家族大教堂在形状上非常相似，但在成因和施工上完全不同。蚁冢的结构和形状自有其理由，但并没有任何白蚁将这些理由表达出来。不会有策划这一结构的白蚁建筑师，也不会有任何白蚁个体知道它们筑巢方式的蛛丝马迹。无需理解亦能胜任。同样，高迪这一杰作的结构和形状也有其理由，但这些主要是高迪的理由。高迪造出这种形状有他的理由；蚁冢成为这种形状也有理由，但白蚁却没有这些理由。树木为何如此分支是有理由的，但这并不是树木的理由。海绵的行为有理由，细菌的行为也有理由，甚至病毒的行为也有理由。但它们并未拥有这些理由，也不需要拥有这些理由。

这些行为背后存在着大量理由，但在一般情况下，生物体并不需要理解这些理由。生物体与生俱来的行为进化过程中良好地设计出来，它们是这些设计的受益者，而无需了解这一过程。这一特征在自然界无处不在，但它往往为我们采用意向立场的倾向所掩盖，我们总会倾向于将行为解释得比实际更理性和有意为之。白蚁是多么巧妙地精心布置蚁冢的通风管道啊！松鼠是多么有远见地储存食物过冬啊！梭鱼是多么狡猾地静候它的猎物啊！这些行为的确是在无情竞争的大自然中取得成功的杰出策略，但其受益者并不需要感谢我们发现了这些策略。我们是表述出这些理由、为这些成功的安排做出解释的首批具有心智的生命。

41

蝉知道素数的含义吗
无需理解的能力之三

为了把数与数字的不同、自由浮动的理据与得到表述的推理的不同搞清楚，让我们看看十七年蝉。1977年，斯蒂芬·杰伊·古尔德写下一篇极富洞察力的、令人钦佩不已的文章，他注意到这样一个古怪的事实：某些蝉的繁殖周期，若以年计，等于素数，例如，十三年或十七年，但从来不会是十五年或十六年。古尔德说：“作为进化论者，我们苦苦寻觅这一现象的原因。为什么会存在这样惊人的同步进化？为什么这种蝉有性繁殖的时间间隔这么长？”答案回想起来很美妙：在这种蝉两次出现之间有一个较长的素数周期，这样，它们便可以降低被其二年生、三年生或五年生的天敌发现并当成一道美餐的可能性。如果这种蝉的繁殖周期是十六年，那么，一年生的天敌就不会针对它们，但是它们会成为二年生或四年生天敌的可靠食物源，而对于与它们步调相合的八年生天敌来说，这将变成一场机会均等的对赌。反过来说，如果它们的繁殖周期并非一个较小数目的倍数，那么，除了那些有幸同它们拥有完全相同或成倍繁殖周期的物种（比如传说中的三十四年生“知了终结者”）会过上养尊处优的生活，它们就不值得被其他捕食者“努力”追踪了。

这种解释的可靠性虽然可能尚未建立，但它并不依赖于假设蝉能理解算术，更别说什么素数了，这一点很清楚。同样，它也不依赖于假设自然选择过程会理解素数。无意识、无理解能力的自然选择过程能够利用某些数的重要性质而根本用不着理解它们。还有一个例子：几何学断言六边形是蜂窝隔间的理想形状，但蜜蜂或大自然母亲并不需要理解几何学。当然，我还可以引用更多的例子来说明进化过程中这种无需理解亦能胜任的数学能力。

42

瞪羚的弹跃
无需理解的能力之四

还记得两个黑盒子的直觉泵所揭示出的令人困惑的规律吗？你不必对盒子有语义解释或意向解释就能知道：α按钮导致红灯亮，β按钮导致绿灯亮。这一显而易见的模式需要一个解释。对于按下按钮的每一种情形，科学家们能精确地了解计算和传输过程的每一步，但是，他们的解释不能普遍化。你需要一个语义解释来说明为什么会存在这种规律。换句话说，“宏观因果”层面上的解释不能“下降”到“微观因果”层面。

对于这种非常普遍的现象，我还可以给出另外一个实例。你可能看过瞪羚被捕食者追逐着穿过草原的视频，也许会注意到其中一些瞪羚在试图逃避捕食者时会高高地跃起。这就是所谓的弹跃。瞪羚为什么会弹跃呢？这显然是有好处的，因为那些很少做出弹跃动作的瞪羚会被捉住吃掉。这是一个已被细致观察过的因果规律，就像按钮与灯之间的规律一样。但这也会让人感到困惑。所有瞪羚以及追逐它们的捕食者的所有细胞中的蛋白质的行为均不能解释这一规律为何存在。

为此，我们需要进化论的一个分支理论，它被称为高成本信号理论（costly signaling theory; Zahavi, 1987; FitzGibon & Fanshawe, 1988）。最强壮、最敏捷的瞪羚跳起来向追逐者宣示它们身强体壮，这实际上释放了一种信号：“别自寻烦恼地追我，我很难抓住哦，专注于我的小伙伴吧，它们不能跳，多容易吃到啊！”捕食者会把这看作是难以伪造的真实信号，进而无视那些做出弹跃动作的瞪羚。这便是自由浮动的理据，瞪羚和捕食者都无需领会这一理由。也就是说，一只能够弹跃的瞪羚完全不必知道弹跃为什么是一个好主意，而狮子等捕食者也用不着理解自己为什么会去捕食相对于那些弹跃的瞪羚而言毫不起眼的猎物。

但如果这一信号不是真实的高成本信号，那么，它便不能维系捕食者和猎物之间在进化过程中的生存竞赛。如果进化过程试图采用一个“廉价”信号，比如甩尾，那么每一只瞪羚，无论强壮还是羸弱，都可以释放

这一信号，它就不会引起狮子的注意，所以，瞪羚不会采用这一信号。根据自由浮动的理据做出的这些解释不能还原到较低的层面，比如分子层面，但重要的是要认识到，即便站在意向立场（从狮子的角度看瞪羚的弹跃行为能得出什么结论）上能够解释瞪羚为什么弹跃，狮子或瞪羚个体也不需要理解弹跃的意义，它们只要近似知道就可以了。

43

世界上第一只哺乳动物是什么
压根儿就不存在

你可能会认为自己是哺乳动物，狗、牛、鲸也是哺乳动物。可是，压根儿就不存在哺乳动物。不可能存在！这里有一段哲学论证可以证明这一点（改述自Sanford, 1975）：

> ① 每个哺乳动物均有一个同为哺乳动物的母亲。
>
> ② 即便存在哺乳动物，它们的数量也是有限的。
>
> ③ 如果曾经存在过一个哺乳动物，那么由第一条可知，一定存在过无数的哺乳动物，这与第二条矛盾，所以不可能存在任何哺乳动物。该说法自相矛盾。

我们完全清楚，哺乳动物是真实存在的，所以，我们只是将其作为一个挑战来认真对待，以便发现其中潜伏着什么样的谬误。这是一个简单的归谬法，因此必须舍弃掉某种东西。我们也大体知道要舍弃掉什么：如果你在任何哺乳动物的谱系树上回溯得足够远，你最终会到达兽孔目，这是一种已经灭绝了的奇怪的过渡物种，介于爬行类和哺乳类之间。

在从明确的爬行类到明确的哺乳类之间逐步过渡的过程中，有着许多难以归类的中间形态。要在逐渐变化的连续谱系上画出一条条界线，我们可以做些什么呢？我们能否确认一只没有哺乳动物母亲的元初哺乳动物（The Prime Mammal），从而否定掉前提（1）？有何依据呢？无论是什么样的依据，都势必要与另一种判定那只动物不是哺乳动物的依据相冲突。毕竟，那只动物的母亲是兽孔目。这不就是证明那只动物属于兽孔目的最好证据吗？假设我们列出了区分哺乳类和兽孔目的10个主要差别，并且说明，拥有5个或5个以上哺乳类标志的动物就是哺乳动物。可为什么是10个而不是6个或20个差别呢？难道不应该给这些差别

按重要程度排序吗？且不提这种判定的武断之处，任意一种一刀切的方法都将产生大量我们不希望看到的结论，因为在从明确的兽孔目动物过渡到明确的哺乳类动物这段非常长的时间内，拥有5个或5个以上的哺乳类标志的哺乳类动物会同只拥有少于5个的哺乳类标志的兽孔目动物交配，产生大量的后代，哺乳类诞下了兽孔目，兽孔目又诞下了哺乳类，如此等等。

当然，我们可能需要一部时间机器才能观察到所有这些异常现象，因为在亿万年后的今天，所有这些细节都已经检测不到了。这倒也无妨，因为这些细节也并不重要。那我们该做些什么呢？我们应该压制划界的欲望。我们不需要划界。我们可以泰然面对这一平淡无奇的事实：所有这些变化在数百万年间逐渐累积，最终产生了无可否认的哺乳动物。同样的道理，我们并不需要标定湖泊、池塘、湿地和沼泽之间的差异，就连湖沼学家也不需要。

然而，哲学家往往有洁癖，他们是挑剔的词语使用者。自苏格拉底坚持要求准确标识出德性、知识、勇气以及类似概念的典型特征以来，哲学家们总是按捺不住这样的想法：通过认定某种可以阻止退行的东西来防止无限倒退，如本例中的元初哺乳动物。这往往让他们得出深陷神秘（或至少是迷惑）之中的学说，也让他们委身本质主义（essentialism）。元初哺乳动物必须是哺乳动物集合中首先拥有全部哺乳动物本质特征的那一只。如果没有可定义的哺乳动物的本质，这些哲学家就遇上了麻烦。然而进化生物学告诉我们，确实不存在这样的本质。所以，作为一个一般规则，我们可以无视哲学家对本质、典型特征或某个“真理制造者”的需求。它通常始于一场徒劳的追寻，可能颇有意思，但最多只能算作无关痛痒的启发。虽然也并非总是如此。

放弃这个需求像是一剂让许多哲学家难以下咽的苦药。用理性的方式从事哲学是从苏格拉底那里传下来的方法，它几乎总是提出“请定义你的用词”这样的要求，从而将大家置于某个本质主义的教条中，即便只是为了论证方便起见。如果我们不得不放弃本质主义，我们所钟爱的一些论证形式就几乎没用了。例如，让我们看看下面这个论证的结构，它始自一个平淡无奇、显然为真的析取式：

① 要么A，要么非A（你如何论证这一点？）。

② 如果你选择路径A，接着，喋喋不休大肆论证一番，那么你将得到结论C。

③ 同样，如果你从非A开始，接着吧啦吧啦论证一番，那你也可以得到结论C！

④ 因此，C成立。

但是，假使明显存在着大量的中间地带，那么对那些说不清它们是A还是非A的情况该如何对待呢？比如哺乳类或非哺乳类、活着的或没有活着的、有意识的或无意识的、信念或非信念、道德或不道德，等等。为了摒弃这种担忧，你必须“划清界限”，区分A与非A，从而驱散所有有关“近似”的谈论。若没有这种泾渭分明的界限，标识出某个本质就从来都是危险之极的，我们也根本无法构造出这种论证。在数学上，这种论证方式所向披靡，因为在这个领域你真的可以划出界限：一个整数不是奇数就是偶数，一个实数要么是有理数要么是无理数，一个多边形要么是三角形要么不是。但在这种抽象的领域之外，这种论证远不够成功。

必定存在着元初哺乳动物，即使我们永远无法知道它存在于何时何地。坚持这一想法是歇斯底里现实主义的一个表现。它使我们认为：只要知道得足够多，我们就必定能看到哺乳动物身上具有某种特殊性质，这种性质可以一劳永逸地定义何为哺乳动物。哲学家有时会说，否认这一点是混淆了形而上学与认识论：混淆了什么是真正存在的和什么是可以知道存在的。对此，我的回应是，思想家确实可能因为混淆了形而上问题和单纯的认识论问题而走上邪路，但这需要证明，而不仅仅是做出断言。

44

物种形成于何时
某个精确但又不为人知的时刻

经自然选择而进化的过程有一个奇怪的特征：它在很大程度上取决于那些“几乎从未”发生过的事件。例如，物种形成，即生物偏离了自己的亲本，从而产生出一个新物种。这是一个极为罕见的事件，但是，已经存在于这个星球上的数以百万计的物种中的每一种都是从某个物种形成事件开始的。每条世系中的每次降生都是一个潜在的物种形成事件，但该事件几乎从未发生过，百万次降生中也不会有一例。DNA的突变也几乎从不发生，一万亿次复制过程中也不会有一例突变，但是，进化却依赖于它。此外，绝大多数的突变要么有害，要么是中性的；偶然的“有利”突变更是几乎从不发生。但进化恰恰依赖于这些极为罕见的事件。

考虑一个关于可能性的直觉泵：就我们所知，目前在这个星球上只有一种人科动物，即智人（Homo Sapiens）。假设五十年后，我们大部分的后代都被某种病毒消灭了，只留下两种幸运的存活者：一千名因纽特人栖居在边远的康沃利斯岛，离格陵兰岛不远，而一千名安达曼人生活在印度洋腹地一个与世隔绝的海岛上。这两个群体相互隔离了数千年，为了适应不同的生存环境各自产生了特有的生理差异，但是，我们并没有理由怀疑他们不属于我们这个物种。假设这些人在地理隔离和生殖隔离的情况下又生存了一万年，最终成为两个物种，繁衍遍布至整个星球。直到有一天他们彼此遭遇，发现没有兴趣同对方交配，少数无意中发生的交配也无果而终。这种由长时间的地理隔离产生的生殖隔离是异域性物种形成的典型标志。

他们也许想知道，这次物种形成事件到底是在什么时候发生的？他们最近的共同祖先很可能生活在三万多年前，但物种形成事件并不是在彼时彼地发生的，但在这两个种群重新相遇之前，物种形成事件仍可能在分离后几千年间的某个时候发生。物种形成事件发生于农业革命之前，还是互联网诞生之后呢？我们不可能心安理得地给出一个笃定的答

案。他们之间一定存在一个最近的共同祖先，道金斯称之为“共同祖先”（concestor），他兴许生活在三万年前，其后代最终有可能在适当的时候成为这两个物种的创始者，但是，我们始终无法确定，这次事件是否开始于那时。

这里应当发生了一个事件，一次对人类以及人类之后的历史举足轻重的降生，这次降生发生在一个精确的时间地点，但直到数千年后它带来的后果被确定了，它才会获得特殊的地位。这从来都是无法预先决定的。倘若有一船或一飞机的岛民来了趟旅行，让两个分支“过早”相聚，这次降生就永远成不了物种形成事件了。可以想象，物种形成事件实际上发生于某个精确但又不为人知的时刻，该时刻介于最初的分离和两个物种最终形成之间，但是怎样才能界定这个时间点呢？根据两个世系间染色体发散（chromosomal divergence）的累积程度，倘若存在跨世系交配的尝试，该时刻很可能就是这类尝试首次无果而终的时刻。但这类有违事实的猜测并没有什么意义。

横贯美洲大陆的铁路将北美野牛分成了诸多生殖隔离的种群，但野牛比尔·科迪（Buffalo Bill Cody）及其同伴很快就将潜在的物种形成事件消灭在萌芽状态，他们几乎将野牛消灭得只剩下一个种群。同种个体常常会因环境事件而分成两个或更多个孤立的群体，在数代间它们保持着生殖隔离，但在绝大多数情况下，这些群体要么重聚，要么有某个群体灭绝。所以，虽然形成物种的第一步必定发生得相当频繁，但它们几乎从不会导致物种形成事件，而真正的物种形成事件需要几百代的时间方能产生出结果。即便你拥有彼时这个世界上每一个分子的物理知识，也绝对无法从首次分离的环境推断出这是不是物种形成事件的开端。“物种”这一概念是一个近似概念。家犬、土狼和狼属于不同的物种，但科伊犬和狼狗的后代却也相当常见，所以，“理论上讲”，我们或许应该将这三类动物仅仅看作不同的品种，而不是犬科下的不同亚种。

我们很难确立起杂交（从不同的亲本物种中产生出后代）的一般规律，想想每个物种的个体之间所呈现出的微妙差异，你便会发现这点毫不奇怪。对此，生物学家并不格外担忧；既然已经理解了创造出所有这些中间形态的过程，他们也就不纠结于定义或本质了。

45

回溯性加冕
人类最近的母系祖先“线粒体夏娃”大约生活在20万年前

一位身在纽约的女人很可能突然就成了寡妇，原因是，在千里之外的道奇城[75]，一颗子弹刚刚击中了某人的头部。“西大荒”那些年，有种人称“寡妇制造者”的左轮手枪。不过，某个场合里的一把“寡妇制造者”是不是名副其实，就连最严苛的犯罪现场调查也无法确定。这个例子从婚姻关系约定俗成的性质中获得了一种跨越时空的奇妙能力：发生在过去的一次历史事件，即一场婚礼，建立起了一种形式上而非因果性的长久利益关系，不管后来夫妻双方是否两地分离、是否发生了不幸，如意外弄丢了婚戒或弄坏了结婚证。

遗传生殖系统隶属于自然而非习俗，它像钟表一样，非常具有系统性，可以让我们从形式上思考绵延几百万年的因果链条，否则，我们几乎不可能命名、指涉或追踪它们。这种系统性让我们可以对那些比婚姻这种形式的关系更加遥远、难以看清全貌的关系发生兴趣，并对其做出严格的推理。和“婚姻”一样，“物种形成”也是一个概念，内嵌于一个严格的、形式上可定义的思想体系，但与“婚姻”不同的是，它不具有那些约定俗成的特点，比如婚礼、婚戒以及结婚证，因而不能通过这些东西加以观察。正如我们刚才看到的，物种形成同样是一个在空间和时间上“远距的”奇怪现象。物种是一个边界模糊、随时间变化的集合，我们只能回溯性地、武断地确认其中某个生物是元初哺乳动物。为了更好地理解物种形成的这一特征，我们可以先来看看回溯性加冕的另一个实例：对“线粒体夏娃（Mitochondrial Eve）”名号的追认，这项追认可不是随意而为的。

单个生物比一个物种有着更为清楚的边界，因此也具有更为明确的身份，即便如此，它们当中依然有大量的中间派。让我们看看最引人注目的例子：有十万亿个左右的细胞拥挤在你的衣服下面，其中九成都不

是人类的细胞！没错，数千种共生访客在数量上超过了宿主细胞，也就是来自你父母结合后形成的受精卵的细胞。这些共生访客中不仅有细菌，还有真核生物，包括单细胞微生物以及多细胞生物：真菌、睫毛上或身体其他什么地方的螨虫、只有用显微镜才能观察到的蠕虫或更大个儿的蠕虫，没准儿还有其他什么东西。你是一个行走的生态系统，有些访客是有害的，比如引起脚癣的真菌、导致口臭或引发感染的细菌，而其他访客却非常重要，如果你把所有的入侵者都赶尽杀绝，那你真会死掉。由于这些共生访客的细胞通常比人体细胞小，就重量而言，你主要是个人类，但它们加起来的重量也不容小觑：大概有好几千克重，也许有4.5千克之多。你身上还有病毒，它们的数量更为庞大。

话说回来，尽管“你”与外界的边界千眼百孔，你还是和其他单个生物一样，很容易就同他人区分开来，有时候，我们可以因特定的作用从进化史中挑选出一个特定的生物。线粒体夏娃就是其中最为有名的一个。线粒体夏娃是一位女性，就母系遗传链而言，她是今天还活着的所有人最近的直系祖先。我们的细胞中都有线粒体，它们仅仅通过母系遗传链传递给我们，所以，在所有今天还活着的人身上，所有细胞中的线粒体直接来自于一个女人的细胞中的线粒体，遗传学家丽贝卡·卡恩（Rebecca Cann）、马克·斯托金（Mark Stoneking）和艾伦·威尔逊（Allan Wilson）称其为“线粒体夏娃”。

线粒体是细胞中的小型细胞器，在新陈代谢过程中发挥着核心作用，从食物中获取身体所有机能所需的能量。线粒体有自己的DNA，准确无误地表明了它们起源于数十亿年前的共生。通过分析今天还活着的各色人种线粒体DNA的模式，科学家已经能够大致推断出线粒体夏娃生活的年代，甚至她生活的地方。初步计算表明，线粒体夏娃生活在大约30万年前的非洲，但在最近，这一数字得到了改进：她生活在非洲，这点几乎可以肯定，而她生活的年代推后到了20万年前。反推出线粒体夏娃生活的年代和地点是一个远比推出其存在更为棘手的任务，对于后者，没有生物学家会怀疑。

先撇开这些争议，考虑一下我们已经知道了有关线粒体夏娃的哪些事情。我们知道，她至少有两个女儿又生了女儿并存活下来，因为如果她只有一个女儿，那么这个女儿将继承线粒体夏娃的冠冕。为了将名字同名号区分开来，我们不妨给线粒体夏娃起名为“艾米”。艾米拥有线粒体夏娃的名号，也就是说，她恰好是今天人类世系的母系始祖。重要的是，除此之外，线粒体夏娃很可能没有什么了不起或特别的地方；她肯

定不是第一个女人，也不是智人的创始人。早期无疑有很多女人属于我们这个物种，但她们碰巧不是所有今天还活着的人的身体内所有线粒体的最近来源。虽然线粒体夏娃有女儿、外孙女，但她很可能并不比她那个年代的其他女人更强壮、更敏捷、更漂亮，或更富生殖力，这一点同样重要。

为了表明线粒体夏娃，也就是艾米可能是多么不起眼，假设在未来的数千代之后，一种致命的新型疾病扩散至整个地球，在短短几年内就消灭了99%的人类。那些对这种病毒有先天抵抗力的幸存者很可能都有着相近的亲缘关系。他们最近的共同直接女性祖先，且叫她贝蒂，可能是生活于艾米千百代之后的某个女人，这时线粒体夏娃的冠冕就会回溯性地传给她。她可能是那次在几个世纪后拯救了我们这个物种的突变的来源，但这对她并没有任何好处，因为取得胜利的病毒在彼时可能尚未获得后来那种致命形态。重点是，线粒体夏娃只能被回溯性地加冕。历史上这个举足轻重的角色并不仅仅由艾米所在时代的一些偶然事件决定，它同样被随后发生的那些偶然事件决定。

谈谈非常严重的偶然事件吧！倘若三岁那年，艾米的叔叔没有救起溺水的艾米，我们所有人，连同我们身上最终来自艾米的线粒体DNA就都将不复存在！倘若艾米的所有外孙女都饿死在襁褓之中，就像那个时代经常发生的那样，我们也将湮没无闻。

同样的逻辑，一定也存在某位亚当：今天还活着的每个男人最近的直接男性祖先。我们可以称他为Y染色体亚当，因为我们所有的Y染色体都是通过父系遗传的，就像我们的线粒体通过母系遗传那样。(76)Y染色体亚当是线粒体夏娃的丈夫或情人吗？绝无可能。父系遗传比母系遗传的时间消耗和能量消耗更少，所以从逻辑上讲，Y染色体亚当生活的时代可能非常晚近，而且他会非常忙于床笫之欢，呃，远胜于埃罗尔·弗林（Errol Flynn）(77)。假使当今最长寿的男性可以活到110岁，那么，逻辑上讲，Y染色体亚当就有可能是他的父亲，不仅如此，这位生活在20世纪初的花花公子还可能是所有较年轻男士的父亲、祖父或曾祖父等。别忘了，男性可以产生无数精子，每次射精就能射出数以亿计的精子，所以，原则上Y染色体亚当用不到一周的时间就能产生出足够的精子从而成为全人类的父亲！然而，通过统计世界各地男性Y染色体当中基因的全部差异，并计算这些突变积累所花费的时间，我们可以估算出，Y染色体亚当生活的年代实际上距今不到10万年。再说一遍，假如某次瘟疫消灭了半数男性人口，Y染色体亚当的冠冕就极有可能加诸一个更为

晚近的先祖。(78)

这里有一个奇妙的事实：每个生物个体，比如你、我、你的爱犬或天竺葵，都可能是某种新物种潜在的创始者，是某种生物的原点，但是可能需要千百代的时间，那种生物才能脱颖而出，被认作新物种，所以，加冕必定发生在很久以后，那时你、我、你的爱犬或天竺葵早已烟消云散。因此，你的父母可能会成为两种人形物种所有成员最近的共同祖先。吃惊吧？吉娃娃犬和大丹犬同属家犬（Canis familiaris），但若人类文明崩溃，它们各自的后代野化，那么，它们就可能比比格犬和巴吉度猎犬更易另立门户：因为没有人类的帮助，大丹犬给吉娃娃犬受精或反过来都是不可能的。然而，跟时间长河中的大多数世系一样，这两个世系极有可能等不到那一天就已灭绝。

据估计，在所有曾经存在过的生物中，有超过90%的生物无后而终。然而，你，连同多年来你数以亿计的先祖，从单细胞生物到爬行类、哺乳类、灵长类，都并不是它们中的一员。你该有多幸运啊！当然，每一根草也有其同样漫长和引以为豪的历史，每只蚊子、每头大象、每株雏菊也都是如此。

46

循环
周期性重复是进化的关键

人人熟知自然界的大循环：昼夜相继；四季更替；蒸发和降水形成的水循环注成湖泊、冲成河流，为这个星球上的每一个生灵提供水分。但并不是所有人都能意识到，从原子尺度到天文尺度的各种循环简直就是隐藏着的高速马达，为所有令人叹为观止的自然现象提供动力。1861年，德国发明家尼古劳斯·奥托（Nikolaus Otto）制成并出售了第一台汽油内燃机；1897年，发明家、物理学家鲁道夫·狄塞尔（Rudolf Diesel）成功发明了柴油机，这两个辉煌的发明改变了世界。不论是四冲程还是二冲程的奥托循环和狄塞尔循环都利用了循环的方式，在完成部分工作后将系统恢复至原初状态，为接下来的工作做好准备。这些循环的细节非常巧妙，它们已由一个有着数百年历史的研发过程所揭示，并且有了进一步的优化。克雷布斯循环是一台更为精巧的超小型引擎，由生物化学家汉斯·克雷布斯（Hans Krebs）于1937年发现，它是进化用了数百万年的时间在生命之初发明出来的。这是一种“八冲程”的化学反应，在对所有生命（无论是细菌还是红杉）至关重要的新陈代谢中将燃料（食物）转变为能量。

克雷布斯循环这类生化循环负责生物界所有的运动、生长、自我修复以及繁衍，周而复始，像一架拥有大量活动部件的钟表装置，其中，每一个钟表都可以重置，都能恢复到再次履行其职能的初始状态。所有这些循环已由生息不止的达尔文大循环做出了优化，历经一代又一代，在漫长的时间中累积着偶然的改进。

在一个完全不同的尺度上，我们的祖先发现了循环的功效，这是人类史前时代最伟大的成就之一：发现重复的制造作用。拿一根木棍同石头摩擦，几乎不会发生任何变化，唯一可见的不过是一些划痕。拿起棍子再做一遍。你还是得不到任何回报。摩擦一百次，也很难看到任何变化。但这样做上几千次，你就能把它磨成一支不同寻常的箭杆。通过不

知不觉的增量累积，循环过程可以创造出全新的东西。这类任务需要将前瞻性和自我控制结合起来，而这种结合本身就是一种新奇的事物，是对其他动物主要发自本能的盲目的重复性建造过程和成型过程的巨大改进。而那种新事物本身也是达尔文循环的产物，通过文化进化这一更为迅捷的循环得到最终的增强，在此过程中，技术复制不是通过基因传给后代，而是通过模仿在无亲缘关系的个体间传播。

第一位祖先将一块石头磨成了漂亮匀称的手斧，在此过程中，他一定看起来非常愚笨：他坐在那里，磨上好几小时都没有明显的效果。但是，在所有单调整重复行为的间隙，潜藏着一个缓慢改进的过程，对此进化设计出的肉眼几乎无法察觉，它们只能看到那些以快得多的速度发生的变化。[79]同样是这种看上去无效的现象，有时也会误导老练的生物学家。在简练明晰的《湿件》（*Wetware*,　2009）一书中，分子细胞生物学家丹尼斯·布雷（Dennis Bray）描述了发生在神经系统的循环：

> 在典型的信号通路中，蛋白质被不断地改性、再改性。激酶和磷酸盐就像一窝蚂蚁那样不停地工作，将磷酸基添加到蛋白质上，又将它们移除。这似乎是一项无意义的工作，而且每一个添加和移除的循环都要花费细胞中的一个ATP分子，也就是一个珍贵的能量单位。实际上，这种循环反应最初被贴上了“无效”的标签。不过，这个形容词是误导人的。将磷酸基添加到蛋白质上是细胞中最常见的一类单一反应，此类反应支持着细胞所要执行的大量计算。这种循环反应远非徒劳，它为细胞提供了必不可少的资源：一个快捷灵动的调节装置。（Bray, 2009, p. 75）

“计算”（computation）一词用得甚为恰当。程序员探索可能的计算空间已有近百年的历史，但到目前为止，他们的发明与发现不过是包含了数以百万计的循环嵌套。事实证明，认知的全部“魔法”就和生命一样，取决于周而复始的循环、“重入”以及自反性的信息转换过程，这些过程广泛地存在于神经元内部的纳米级生化循环、感知系统中具有预测编码能力的生成与测试循环（Clark, 2013），以及全脑睡眠循环，即脑电图记录显示出的大脑激活和复原的剧烈起伏。生命中每一处改进的秘诀总是千篇一律：练习、练习、再练习。

要记取的是，达尔文式进化只是一种累积性的改善循环。当然还有其他类型的循环。智能设计鼓吹者之类的人会争辩说，既然经自然选择

而进化取决于繁殖，那么，达尔文式的方案就无法解释第一个生物、第一个具有繁殖能力的生物是怎么来的。如此一来，生命的起源问题似乎变得不可解了，有着“不可还原的复杂性”（Behe, 1996）。生命的起源肯定极度复杂、有着精妙的设计，它一定是个奇迹。神创论者希望我们把生命之前无生殖力的世界想象成一团混沌，充满了四散飘零的化学物质的碎片，再由一场风暴把它们重新组装成那架人人尽知的飞机。倘若我们陷入这样的迷思，生命的起源问题就会看起来令人生畏，愈发难解。

但是，如果我们提醒自己，进化的关键过程是周期性的重复，基因复制只是其中一个得到高度改进和优化的实例，我们就有可能将奥秘转化为一道智力谜题：四季更替、水循环、地质循环和化学循环，所有这些循环是如何历经数百万年的回转，逐渐积累起了开创生物循环的先决条件的？也许，刚开始的一千次“尝试”都毫无效果、功亏一篑。但就像乔治·格什温和巴迪·迪斯尔瓦（Buddy DeSylva）在那首性感的歌曲里提醒我们的，看看你“再做一次”、再一次、再一次……会发生什么。(80)

不错，当你面对生命世界和精神世界中的“魔法”时，一个好用的经验方法便是去探寻那些承担着全部艰辛工作的循环。

47

青蛙的眼睛究竟将什么告诉了青蛙的大脑

扩展适应的最好例子

J. Y.莱特文（J. Y. Lettvin）及其同事合著的那篇有名的论文《青蛙的眼睛究竟将什么告诉了青蛙的大脑》（*What the Frog's Eye Tells the Frog's Brain*，1959）是认知科学的早期经典文献之一。这篇文章表明，青蛙的视觉系统对视网膜上移动的小黑点非常敏感，在几乎所有的自然环境中，小黑点都是附近飞虫投下的阴影。这个“飞虫探测器”恰当地与青蛙舌头上的微力触发器相连接，轻松地解释了青蛙如何在一个残酷的世界中生存繁衍。那么，青蛙的眼睛究竟将什么东西告诉了它的大脑呢？飞虫？还是“飞虫或假飞虫”？还是某种K型物件，也就是能可靠地触发视觉小装置的某种东西（回想下双币机）？

达尔文派的意义理论家，比如露丝·米利肯、戴维·伊斯雷尔（David Israel）还有我，曾探讨过这个例子，而进化论的坚定批评者杰里·福多尔也揪住这一例子不放。他认为，此类意义的进化式阐论在他看来都是错的：这些意义太不确定了。它们本该对“这儿有飞虫”和“这儿有飞虫或假飞虫”这类蛙眼报告做出区分，但它们没有。不过，福多尔是错的。只要我们能够确定青蛙的选择环境，就能用它来区分各个意义选项。要做到这一点，我们可以采纳与之前解决双币机状态的意义问题时完全相同的考虑。只要在那个选择环境下可以单独挑出一类特定的场合，其中不存在别的东西，那么蛙眼报告真正意味着什么就根本不重要。将青蛙送往巴拿马，或更准确地说，送往一个全新的选择环境，我们就能清晰地认识到这一点。

设想科学家从某种濒临灭绝的食虫蛙那里采集到一个小型蛙群，把它们养在一处新的环境里：这是一座特别的蛙园，没有飞虫，只有饲养员定期地在它们周围抛撒饲丸。让饲养员感到高兴的是，系统运作正常；青蛙可以伸出舌头吃到饲丸，茁壮地成长起来。一段时间后，出现了一群幼蛙，它们从未见过飞虫，只吃过饲丸。它们的眼睛将什么东西

告诉了它们的大脑？倘若你坚持认为意义并未改变，你可就难了。该例是人为构造出来的，是对一直以来发生在自然选择过程中的扩展适应（exaptation）的清楚例示。扩展适应是指重新利用已有结构获得一项新功能。正如达尔文谨慎提醒我们的，为着新目的重新利用已有机制是大自然成功的秘诀之一。对于那些希望得到解说的人，我们可以把事情讲得更透彻。假设饲养的青蛙并非完全一样：青蛙视觉系统的饲丸探测力发生了变异，导致一些青蛙吃到的饲丸较少，自然它们留下的后代也就比较少。于是，短期内就会出现对饲丸探测的选择，无可否认会出现这种选择，虽然这种选择到底从什么时候起才算出现是无法确定的。千万别指望某个精确的时刻可以指示蛙眼所看到的东西正好发生了变化。根本就不存在元初哺乳动物，也根本不存在什么元初饲丸探测。

倘若该触发条件下各个蛙眼不发生“无意义”或“不确定”的变异，那么就不会有为新目的而进行选择所需的原材料了，该原材料正是盲目变异。福多尔等人将不确定性视作意义进化式阐论的缺陷，实际上，它不过是所有此类进化的先决条件。有想法认为，一定存在某种确定了蛙眼真正意指的东西：某个可能不为人知、根植于蛙性的命题，它精确地表述了蛙眼将什么东西告诉了它的大脑。这种想法恰是本质主义的意义观或功能观。意义，和它直接依赖的功能一样，一开始并不是某种确定的东西。它的产生不是来自嬗变（设计空间中的巨大飞跃）或特殊的创造，而是来自周遭环境中往往是渐进式的微小变化。

48

穿越巴别图书馆
福尔摩斯的高效推理

杰出的天文学史家奥托·诺伊格鲍尔（Otto Neugebauer）在1988年收到一张希腊草纸残篇的照片，上面记着一列数字。发件人是一名古典学家，他对这段残篇毫无头绪，想知道诺伊格鲍尔会有什么想法。这位89岁的学者算了下每行数字间的差值，确定出它们的最大值和最小值，由此断定这张草纸一定转译自巴比伦楔形文字泥板书“G列”的一部分，记录的是巴比伦阴历的“B系统”！这是一种用来计算某一时段内天体位置的星历。诺伊格鲍尔是如何做出这个福尔摩斯式的推理的呢？基本演绎法如下：诺伊格鲍尔发现这段希腊文是一串六十进制而非十进制的数列，正是巴比伦人对月球位置精确计算的一部分，也就是G列。计算星历有很多种不同的方法，但诺伊格鲍尔知道，任何独立制作历书的人都有自己的一套系统，他们得出的数字不会完全相同，即便这些数字可能非常接近。巴比伦的B系统极其优秀，所以，该设计方案连同它的每一个细节经过翻译都让人欣慰地保存下来了（Neugebauer, 1989）。

诺伊格鲍尔是一位伟大的学者，但你也可以追随他的脚步，做出一次类似的非凡推理。假设你收到下面这段文字的复印件，并被问及同样的问题：它是什么意思？来自哪里？

Freunde, Römer, Mitbürger, gebt mir Gehör! Ich komme, Cäsars Leiche zu bestatten, nicht, ihn zu loben.

在继续阅读之前，试试破译它吧。即便你不懂古德文的哥特体，甚至不懂德语，你还是能将它破译出来。再仔细看看。尝试大声读出这些单词，不要担心发音问题。明白了吗？多么令人印象深刻！诺伊格鲍尔可以鉴定出巴比伦星历G列，你不也很快地认出这个片段一定是伊丽莎

白时代某个悲剧的德语翻译的一部分吗？准确地说，是第3幕，第2场，79～80行。一旦想到它，你就知道它绝不是其他东西！在其他任何情况下，这段连缀成句的德文字母序列出现的可能性都是非常微小的。(81)

参与创造出这个序列的研发过程过于特殊，很难随随便便就复制出来。为什么呢？是什么东西标志着这样一串符号的特别之处呢？心理学家尼古拉斯·汉弗莱（Nicholas Humphrey, 1987）以一种更为戏剧化的方式让该问题变得生动起来：如果你不得不把下面的一件伟大作品“置之脑后”，你会选择哪件？牛顿的《数学原理》、乔叟的《坎特伯雷故事集》、莫扎特的《唐璜》，还是埃菲尔铁塔？汉弗莱的回答如下。

> 要是不得不做出选择的话，我丝毫不怀疑，答案应该是《数学原理》。怎么会这样？因为在这些作品中，牛顿的作品是唯一可被替代的。理由很简单：如果牛顿没有写它，自然会有其他人写，也许相差不了几年的时间……《数学原理》是人类智慧的一座丰碑，而埃菲尔铁塔是一项罗曼蒂克工程，相比之下没那么伟大。然而，事实是，埃菲尔的成就出自自家之手，而牛顿的成就仅仅是神来之笔。

49

谁是《帕姆雷特》的作者
“延伸的表现型”为你揭开谜底

设想弗兰肯斯坦博士设计并制造了一个怪物帕士比亚，它随即坐起身来，写下了一出名为《帕姆雷特》的戏剧。那么谁是《帕姆雷特》的作者呢？

首先，请读者们注意与这个直觉泵不相干的东西：我既没说帕士比亚是一个硅铁机器人，也没说它就像原版弗兰肯斯坦的怪物一样，由人类的有机组织构成，或者通过纳米工程由细胞、蛋白质、氨基酸或碳原子构成。只要弗兰肯斯坦博士造出了这个玩意，那它是用什么材料制成的根本就不重要。要想造出一个既小巧敏捷又节能高效，还能坐在凳子上写剧本的机器人，唯一的办法似乎是先着手制造人造细胞，然后将那些美妙精巧的马达蛋白或其他一些碳基纳米机器人置于其中。这些科学技术问题虽然都很有意思，但它们不是我们的关切所在。出于完全相同的理由，如果帕士比亚是硅铁机器人，内置程序高度复杂，那它可能比一个星系还要庞大；我们可能不得不打破光速的限制，否则，在人类有生之年我们都见不到它。这些技术上的限制通常被列为此类直觉泵的禁区，不过因为它们没有任何实质性的影响，所以这回我们可以打打擦边球。转动直觉泵上的旋钮，看看是不是这样。如果弗兰肯斯坦博士想用蛋白质之类的东西制作他的人工智能机器人，那是他的事。他的机器人若能和正常人类交配，并因此产生出新物种，那可就太神奇了。不过我们关注的是帕士比亚的心血：《帕姆雷特》。回到我们的问题：谁是《帕姆雷特》的作者？

要把握住这一问题，我们必须到帕士比亚的内部一探究竟。一个极端的情况是，我们发现，它自带存储器并内置了一份文件或《帕姆雷特》的基础存储版，一切都已加载就绪。在这种极端情况下，弗兰肯斯坦博士肯定是《帕姆雷特》的作者，而帕士比亚只是一部存储和递送装置，一台格外高端的文字处理器。所有研发工作早在复制给帕士比亚之

前就已完成。

想象《帕姆雷特》及其浩如繁星的邻居都位于巴别图书馆，如此一来，我们就能更加清楚形象地认识到整个过程。《帕姆雷特》是如何到那儿的？其研发轨迹是怎样的？如果我们发现在帕士比亚的记忆被构建出来并被充满信息的时候，整个旅程已经结束了，那么，我们就知道帕士比亚并没有在搜索过程中发挥作用。往回推，若我们发现，帕士比亚的唯一作用是先通过拼写检查器运行已存有的文本，再利用它引导打字，我们也不会认可帕士比亚的作者身份。拼写检查是整个研发工作中，可以测量但极为微渺的一部分。

《帕姆雷特》的孪生文本则构成了一个相当大的星系：大约有一亿册不同的突变文本，每一本中仅有一处未订正的排印错误；如果把每页上有一处排印错误的文本也算进来，我们就得面对数量浩瀚的变异文本。再往回推点儿，一旦从排印错误（typos）过渡到思维错误（thinkos）[82]，也就是那些可以说成是错误或次优选择的字眼，我们就得严肃考虑作者的身份了，这与单纯的复制编辑工作形成了鲜明的对照。复制编辑工作相对琐碎，尽管它对作品的最终成型有着不容忽视的重要作用，但其重要性能够很好地体现在设计空间之中：每个微小的提升都有其价值，甚至有时一点点提升都会将你带至一条全新的轨迹。和往常一样，说到这里我们引下现代主义建筑大师路德维希·密斯·凡德罗（Ludwig Mies van der Rohe）的名言："上帝存在于细节之中。"

现在转动直觉泵上的旋钮，看看另一个极端：弗兰肯斯坦博士把大部分工作都丢给了帕士比亚。最现实的情景肯定是弗兰肯斯坦博士为帕士比亚装上了一套虚拟的过去：整整一生的伪记忆。在对弗兰肯斯坦博士设置的编剧欲望做出回应时，帕士比亚便可以利用这套记忆。可以设想，这套伪记忆里不仅有在剧院的许多个夜晚、阅读书籍，还有不求回报的爱、骇人听闻的死里逃生、可耻的背叛等。现在会发生什么呢？也许，网上某个"人情味的"故事片段会刺激到帕士比亚，使其进入狂乱的生成与测试过程：它在记忆中到处翻找可用的花边新闻和主题，将找到的东西变形，然后杂糅成希望满满、力争完善的临时结构，其中大部分的结构会被侵蚀性的批评过程逐步分解，不过，批评过程有时也会发掘出零星有用的东西，诸如此类。这个多层级搜索过程的每一层都或多或少地受诸多多层级的、内部生成的评价引导，评价函数负责对持续进行的搜索得出的各种结果的评价（的评价）做出响应，然后再有对评价函数的评价、对函数评价的评价的评价（循环套循环套循环……）做出响

应。

现在，假使神奇的弗兰肯斯坦博士预见到了这一活动的所有细节，直至最混乱无序的底层，并人工构建了帕士比亚的虚拟过去及其所有的搜索机制，仅仅是为了产生《帕姆雷特》这部作品，那他就还是《帕姆雷特》的作者，甚至就是上帝。

如此浩瀚的事前知识根本就是个奇迹。为幻想恢复一点点现实色彩吧，我们可以将旋钮转到不那么极端的位置：假设弗兰肯斯坦博士无法预见到所有细节，他把完成设计空间中的一部文学作品或其他什么东西的轨迹这一艰苦工作的绝大部分交给了帕士比亚，这项工作稍后会由帕士比亚内部进行研发确定。通过如此简单地转动旋钮，我们已经逼近了现实，因为我们已经有了令人印象深刻的人工智能作者，它们大大超出了其创造者的预见。

到目前为止，尚未有人创造出一个值得认真对待的人工智能剧作家，但IBM的人工智能棋手“深蓝”、戴维·柯普（David Cope）设计的人工智能作曲家EMI所取得的成绩，在某些方面，堪比人类创造力的最高成就。

是谁击败了加里·卡斯帕罗夫（Garry Kasparow）这位世界国际象棋卫冕冠军？既非计算机科学家默里·坎贝尔（Murray Campbell），也非其团队里的任何人，是“深蓝”打败了卡斯帕罗夫！在象棋比赛中，“深蓝”能比他们任何人更好地布局。他们当中没人能在对阵卡斯帕罗夫时赢得比赛，但“深蓝”可以。是的，不过……没错，但是……你可能会认为，“深蓝”是赢了卡斯帕罗夫，但它的穷举搜索法相比卡斯帕罗夫的思考，是完全不同的探索过程。

然而，事实根本并非如此。至少在想要改变这种对创造性的达尔文式解释时，并非如此。卡斯帕罗夫的大脑由有机物构成，它同“深蓝”的构造差异非常大，但就我们所知，到目前为止，它依然是一种大规模并行的搜索引擎，随着时间的推移，它建立了一系列出色的启发式“修剪”技术，得以避免把时间浪费在不可能的分支上。毫无疑问，研发上的投入在这两种情况下有不同的面貌：卡斯帕罗夫有一套方法可以从以前的棋局中提炼出良好的构思原则，使他能够忽略大部分的游戏空间，而“深蓝”却必须不厌其烦地遍历整个空间。卡斯帕罗夫的“洞察力”极大地改变了其搜索路径的形状，但这并不构成“一种全然不同的”创造方法。“深蓝”的穷举搜索法可以通过一些算法将某种类型的搜索路径忽略，这项任务虽然困难，但并非不可能。一旦识别出这些可忽略的搜索

路径，它就可以就像卡斯帕罗夫那样在恰当的时候重复利用这些研发过程。“深蓝”的设计者做了大量此类的分析工作，将这种禀赋赋予了“深蓝”，但卡斯帕罗夫也同样受惠于棋手、教练、国际象棋专著的丰富成果，这些成果被安装在了他的大脑习惯里。

就这一点而言，美国国际象棋棋手博比·菲舍尔（Bobby Fischer）曾提出了一个建议，仔细想来会很有趣。菲舍尔希望恢复国际象棋比赛中理性的纯洁性，要求在开局时白方的强子[83]随机摆放在后排，黑方强子按照同一随机顺序镜像摆放，但要确保每一方都有一个白格和一个黑格放“象”，并且保证“王”放在两个“车”之间。该建议一旦被采纳实施，马上会使大量死记硬背的开局完全废掉，因为这些开局出现的可能性会变得微乎其微，这对人类和机器棋手来说是一样的。人们又得被迫去依靠那些基本原则，并且不得不实时做出更多艰辛的设计工作。这种规则上的变动是否对人类更有利，现在还很不明朗，这完全取决于人类棋手和机器棋手相比，哪种棋手更倚重于机械记忆，即依赖于先前探索者的研发，只进行最低限度的理解。

事实上，即便对于“深蓝”来说，象棋的搜索空间也太大了，很难做到详尽无遗的实时探索，所以，像卡斯帕罗夫一样，它需要冒着可预期的风险修剪搜索树，同样，它也常常需要预先估算这些风险。在人类和机器这两种非常不同的构造中，都存在着大量的“蛮力”计算。毕竟，神经元不知道下棋是怎么回事。它们所做的任何工作都必然是某种蛮力式的工作。

为了支持计算，我似乎在循环论证，将卡斯帕罗夫的大脑描述成是按AI的方式在工作，但工作总得去做，而且并没有人明确提出任何其他的工作方式。说卡斯帕罗夫动用了“洞察力”或“直觉”毫无意义，因为这么说只是意味着卡斯帕罗夫本人没有特殊渠道，无法洞见到自己如何发现了一个好的处理方案。所以，既然包括卡斯帕罗夫本人在内，没有人知道卡斯帕罗夫的大脑是怎么工作的，那么，说卡斯帕罗夫的方式“完全不同于”“深蓝”的探索方式就没有任何证据。当有人“想当然地”坚称卡斯帕罗夫的方式非常不同时，他们应该记取这一点。到底是什么引发人们冒险提出这样一种观点的呢？是一厢情愿？抑或是出于恐惧？

你会说那不过是下下棋，又不是艺术。国际象棋与艺术相比微不足道，因为国际象棋的世界冠军现在是一部电脑了。这下轮到作曲家和计算机黑客柯普的音乐智能实验（Experiments in Musical Intelligence，简称EMI，2000，2001）登场了。柯普打算创造出一套纯粹的提高效率

的装置，一架作曲机器，用来帮助他突破任何一个作曲家都会面临的创作瓶颈，它利用高新技术对钢琴、曲谱、录音机等传统搜索方式进行了扩展。

随着EMI能力的提高，它合并了越来越多的生成与测试过程，将自己变成了一个真正的作曲家。当为EMI输入巴赫的音乐时，它会创作出巴赫风格的音乐作品。当输入莫扎特、舒伯特、普契尼或斯科特·乔普林的作品时，它能轻而易举地对这些作品的风格做出分析，然后生成相应风格的新作品，比柯普本人，甚至比几乎所有人类作曲家的模仿之作更胜一筹。当输入两个不同作曲家的作品时，EMI能立即生成将两者的风格怪异地结合在一起的曲子；当同时输入所有这些不同风格的作品时（你可能会说这没有品位），它会基于所有这些作品的音乐“体验”来创作。而且它输出的结果还可以反反复复多次输入，并掺杂以任意MIDI格式[(84)]的音乐，结果，EMI形成了它的“个人”音乐风格，这种风格清晰地显示出它受惠于大师们的作品，而同时它又毫无疑问地对所有这些“体验”做出了独特的整合。现在，EMI不但可以写出二声部创意曲和艺术歌曲，还能够创作出完整的交响曲。据我所知，它已经创作了一千多部交响曲。这些作品的品质足以打动作曲家、音乐教授等专业人士，对此，我可以现身说法：它创作的普契尼式咏叹调曾令我几度哽咽。柯普不能自称他是EMI创作的交响曲、赞美诗以及艺术歌曲的作者，正如坎贝尔不能声称是他击败了卡斯帕罗夫一样。

在一位达尔文主义者看来，级联起重机中的这一新元素不过是漫长历史中的崭新一页，我们应该认识到，作者与其作品之间的边界和级联中的其他所有边界一样具有渗透性。理查德·道金斯（1982）注意到，河狸的坝同它的牙齿和毛皮一样，均是河狸的表现型，或者说是延伸的表现型，这为进一步观察到人类作者的边界也完全服从于同样的延伸提供了基础。事实上，几个世纪以前我们就知道了这一点，我们制定了各种半稳定约定（semi-stable convention），用它们来处理鲁本斯、鲁本斯画室或鲁本斯学生的作品。[(85)]只要有协助方存在，我们都可以问问：谁在帮助谁，何为创造者，何为被造者。

50

虚拟旅馆的噪音

妙用“意外”是创新的关键

考虑下虚拟世界和现实世界的差异。如果你想建一家真实的旅馆，就得花费大量的时间、精力和材料来做相邻客房的隔音。如果是一座虚拟旅馆，你就完全不必这么做。对于虚拟旅馆，若想让客人听到隔壁的声音，你反而得加上那种功能。你不得不加入非隔离性，还得加入阴影、气味、震动、污垢、脚印以及磨损。所有这些非功能性的特征在现实世界中本来就有。这些让虚拟世界看上去更为真实的东西有个统称：碰撞检测（collision detection）。如果曾经制作过电脑游戏，你很快就会意识到，仅仅在屏幕上放上一个运动着的形体还远远不够。这些形体会直接穿过彼此而不产生任何影响，你必须为更新循环加入碰撞检测机制，该循环会时刻干预程序中对象的行为，并使得对象不断询问自己：“我是不是碰上了其他东西？”

在《马罗之挽歌》（*Le Ton Beau de Marot*, 1997）一书中，侯世达讨论了创造过程中“自发侵入”（spontaneous intrusion）的作用。在现实世界中，几乎所有的东西都会留下痕迹、造成阴影、散发气味、发出声音，它们为自发侵入提供了大量的机会。这也恰恰是虚拟世界缺乏的东西。从计算机建模者的角度看，虚拟世界的一大好处就是寂静：除了你提供给它的东西，什么也不会发生。这就让你可以从一张白板开始为模型逐个加入特征，从而得到一个拥有所需效果的最简模型。

噪声的缺席使计算机在模拟进化方面受到了极大限制，因为经自然选择的进化依赖于噪声，它能将意外发生的噪声转化为信号，从而化腐朽为神奇。计算机图形艺术家卡尔·西姆斯（Karl Sims）的《虚拟生物的进化》（*Evolved Virtual Creatures*, 1994）是早期计算机模拟进化过程的佼佼者，迄今为止，它仍然是最令人印象深刻的模型之一。

西姆斯随机装配了一些带关节的虚拟物体，它们的虚拟肌肉可以使这些关节活动，接着，他让这些物体以虚拟环境中的物理规律在其环境

中进化。程序会自动选择那些在实验中移动得最远的装配方式，让具有这种装配方式的物体虚拟交配产生后代，再在后代中重复这一实验。在没有任何智能设计者干预的情况下，这些活动物体进化出了越来越出色的游泳、行走以及跳跃能力。这些设计远非随机，它们表明，虚拟进化能多么有效地近似发现优秀的设计原则，并近似重塑自然中那些令人惊叹的多样特征。

这是一个让人叹为观止的演示，它向我们展示了一个相对简单的模型中能够包含多少东西，但它也表明了虚拟进化受到的限制。西姆斯设计的这个简单的发展系统以全部基因组作为输入，输出新生物体，但这一过程在后台运行，并不是虚拟世界的一部分。结果，一些由虚拟碎片引起的意外颠簸、震动和碰撞既不能缩短也不能延长基因组，更不能改变基因表达的规律。所有的机制并非存在于这个虚拟世界及其中的虚拟物体上，因此它是不变的。比如说，西姆斯创造的这些东西就根本没法进化出一条新的染色体。整个遗传系统位于模型之外，并未直面自然选择，仅仅通过设定的规则传递遗传信息。有关这一现象的其他例子见第51章。

在有关创造性的计算机模型中，你应该会遇到一些躲不开的垃圾和噪声。隔壁房间小杂音的自发侵入可以悄然改进这些过程，也许只是意外，也许在某种程度上具有破坏性，但无论如何它开辟了新的可能性。无论对新的基因组、新的行为，还是新的旋律来说，利用意外均是创新的关键。

让我澄清一下我没说的是什么。西姆斯的问题不在于他所创造的这些能够进化的东西并非由碳或蛋白质、血红蛋白这类物质构成。问题是，它们是虚拟的。它们生活的虚拟世界比生物进化的世界简单得多。柯普的EMI也面临同样的问题。尽管该软件非常奇妙，但它还是比人类的音乐世界要简单上好几个数量级。令人愉快的是，这两个例子表明了你从一个一尘不染、寂静无声、高度抽象的世界那里能得到多少东西。

可以设想，通过添加越来越多的垃圾、越来越多的碰撞机会，我们就可以对柯普的EMI、西姆斯的工作或其他的任何人工生命、人工造物做出改进。这将给他们的虚拟世界提供更多的虚拟物件，而且你永远都无法知道会有什么巧事发生。但是，考虑下这类建议多么有违直觉：

> 无论你在为什么东西建模，确保每一个现象、每一个子程序、在那个世界发生的一切，都能在世界上散播各种非功能性

的影响：产生一些不和谐的噪声、留下一些痕迹、散布一些灰尘、引发一些震动等。

为什么呢？加入这些噪声为的是什么？加入它们并不是为了什么，加入噪声只是让其他所有进程可以将其作为一个潜在的信号源，并通过创造性算法的炼金术使其有可能转变成某种东西，例如某种功能、某个艺术品或意义。宇宙中的每一个设计增量都始于某个机缘巧合的时刻：回过头去看，两条轨迹的意外交叉会产生出某种东西，而不仅是发生了一次单纯的碰撞。但是，如果计算机建模者过于遵循这条建议，他们就浪费了效率，而效率正是计算机伟大之所在。因此这里有一种微妙的自平衡。不难发现，计算机建模表现出的创造力面临收益递减，这并没有什么神秘的原因：为了一步步地逼近人类的创造力，模型必须变得越来越具体，必须越来越多地为只有实体间才能发生的偶然碰撞建立模型。

51

哈尔是赫布和爱丽丝的孩子吗 克雷格·文特尔说：是

已故的伟大进化论理论家乔治·威廉姆斯（George Williams）坚持认为，将基因等同于DNA分子是个天大的错误。这就好比说《哈姆雷特》就是纸上的一连串墨迹一样。当然，任何莎士比亚戏剧的复制品一定是由某种东西构成的，即便不是墨迹，也会有其他形式的载体，比如，电脑屏幕上的字符，或刻录在CD上的二进制代码串，但是，戏剧是抽象的、载有信息的，可以更换媒介存储。根据这一思路，基因作为合成蛋白质的配方，同样也是抽象的、载有信息的。我一直认为这样的思路是正确的。但是，对此向来有人反对，他们认为这样来设想基因没有意义。针对他们，特别是生物哲学家彼得·戈弗雷-史密斯（Peter Godfrey-Smith），我构想了一个小巧的直觉泵。

赫布和爱丽丝想要一个孩子，他们做了以下事情。

1．他们做了基因组测序，随后收到一封关于其基因组的数据文件，这份文件包含两行有着30亿个字母的序列：A，C，G，T……

2．然后，他们为自己的基因组编写了一小段程序，该程序采用细胞减数分裂算法，能够随机产生虚拟的精子和卵子，随后，这两类虚拟配子随机地经计算机模拟结合，于是，产生了一个新基因组的详述文件。经法医鉴定，它描述的确实是赫布和爱丽丝的后代所拥有的DNA。到目前为止，所有这一切纯粹是一个计算机字符串的重写过程。

3．通过不断的克隆，这个详述文件被用来表达整个基因组的DNA，A代表腺膘呤，C代表胞嘧啶，G代表鸟嘌呤，T代表胸腺嘧啶。文特尔的实验室现在就能实现这个过程。

4．接着，这个基因组将被植入某个人类卵子的细胞核

> 中。由于在移植之前该卵子的DNA就被去掉了，因此它是谁的卵子并不重要。该受精卵通过某种常见的方式最终变成了一个“试管婴儿”。
>
> 那么，最后得到的这个婴儿哈尔，是赫布和爱丽丝的孩子吗？对我而言，哈尔无疑是他俩在生物学意义上的后代，因为他用到了赫布和爱丽丝的全部遗传信息，如果哈尔是以正常方式降生的，这些遗传信息的应用方式也不会有任何差别。这个直觉泵突出强调了繁殖过程中的重点：信息，以及信息的因果传递。在本例中，遗传信息以“A”“C”“G”“T”的ASCII码形式传递，而不是分子形式。这条因果链也满可以通过通信卫星传递，而不是通过更为直接的生化过程传递。（改自我2010年4月26日的私人通信）

戈弗雷-史密斯也同意哈尔是赫布和爱丽丝的后代，但是，他对我的解释有所保留（Dennett, 2010; Godfrey-Smith, 2011）。本着同样具有建设性的批判精神，我承认，这对夫妇生育子女的方式和我们通常的生育方式之间有一个生物学上的重大差别。如果每个人都像这对夫妇这么做，那会怎样呢？由于来自赫布方面的遗传信息并没有像通常那样搭乘精子进入卵子，所以精子的活性将不再承受自然选择的压力，在其他条件不变的情况下，精子活性将在后代中下降，毕竟用进废退啊。不过，我依然坚持认为这个直觉泵清楚地表明：遗传信息的分子结构之所以能够在许多世代间大致保持不变，是因为它具体体现了它载有的遗传信息。

假使存在某种可以用来保存信息的替代结构，进化就会畅通无阻地持续下去。这种说法可以在另一个直觉泵中予以深入研究：设想在另一个星球上，“奇数代”使用A、C、G、T这种DNA序列，而“偶数代”则使用完全不同的双螺旋结构，就叫它XNA吧，XNA由P、Q、R、S构成，拥有不同的分子结构。可以假设，子代的XNA分子是以父代的DNA为模版通过某种类似于信使RNA的机制生成的，这种机制可以在两种不同的生化语言间进行“翻译”。在第三代中，这些信息将通过另外一种信使机制翻译回去，以此类推。因此，你若想要孩子，你伴侣的基因就必须与你的基因用同一种语言写成；不过，你后代的基因组就会用另外一种语言写成。

俄狄浦斯式的结合将无法生育，还可能会出现许许多多罗密欧与朱

丽叶式的悲剧，恋人们因为来自不同的基因群体而无法生育。他们也可以接受这种没有结果的性生活，领养孩子，甚至接受卵子或精子捐献，抚养只有一半亲缘关系的孩子。在这样一个世界里，撇开这些社会并发症不论，进化还是会一如既往地通过世代间不同的编码系统传递适应性的遗传信息，也会继续传递遗传病等。如果你乐意，可以设想这两种编码系统的构造完全不同。相同的基因，不同的分子结构。每个基因有两种形式，其差别就好比“cat”与“chat”（法语“猫”）、“house”与“maison”（法语“房屋”）之间的差别。注意该例与两个黑盒子的相似处：两个载体在句法上、结构上不相同，但它们拥有相同的信息，也就是相同的语义。

52

模因

不论好与坏，它都是我们人类的共生体

我尚未提及模因（memes），这可能让一些读者担忧我是不是把它们给忘了。怎么会呢！模因概念是我最钟爱的思考工具之一，关于它我有说不完的话，这一整本书可能都放不全呢！关于模因，我在其他地方已经说过很多了（Dennett, 1990, 1991a, 1995a）。出于很多理由，人们几乎发自本能地不喜欢这个概念，所以他们会被一些批评意见吸引。因此，我觉得我有必要再说说模因，并回应一下所有那些严肃的批评人士和神经兮兮的模因厌恶者，不过这恐怕需要专门写一本小书了。在此期间，那些想深入了解模因的可以看看我的文章《新复制体》（*The New Replicators*, 2002；以及Dennett, 2006a）。

不过，作为预览，下面先就模因概念给出一段简介。这里所指的是严肃的模因概念，而不是网民们过分流行、不甚严格的用法。正如理查德·道金斯（1976）指出的，当他将模因概念作为一个能够自行复制的文化项引入时，生物学的基本原则是，所有生命都凭借自复制体的差异性生存而进化。

> 基因，即DNA分子，恰好就是这种自复制体，它们广泛地存在于我们这个星球上。或许还有其他类型的自复制体。如果有的话，只要能满足其他一些特定条件，它们就几乎无可避免地会成为某种进化过程的基础。
>
> 但是，我们非得去遥远的世界方能找到其他类型的自我复制，以及随之发生的其他类型的进化吗？我认为，就在我们这个星球上，已经出现了一种全新的自复制者。它就在我们面前。它尚处在婴儿期，仍然在它的原始汤里笨拙地漂流着，但是，它产生进化改变的速度，将古老的基因远远地甩在了后面。（Dawkins, 1976, p. 206）

在思考人类的文化和创造力之际，这一思考工具流露出的两个主要的见解极大地改变了我们的思想。首先，模因粉碎了一个诱人的想法，即要想达成一个出色的设计只有两条路径：要么是基因的，要么是天才般的。在模因打开他们的视野之前，大多数思考者认为，人类生命中的某种东西若展现出足以达到目的或实现高效功能的适应性迹象，它要么是基因自然选择的产物，要么是人类慎思、理解、带有意图的思考的产物，即智能设计。奥格尔第二定律，即“进化比你聪明”似乎将这两个选项奉若圭臬，但事实上还有第三种随处可见的可能性：非遗传的文化选择。自然选择既赋予我们基因，又在同一过程中实现了这种非遗传的文化选择。人们在一百多年前对布列塔尼人的独木舟的观察为我们提供了一个鲜活的实例：“每条小舟都仿制自其他小舟……是大海造就了这些小舟，她选择了有用的形式，毁掉了其他形式。”（Alain, 1908）显而易见，这是自然选择。岛民的规则很简单：哪艘船能从大海中安然返航，就仿制哪艘船！他们兴许考虑过造船学原理，这些考虑也许事后确认了他们青睐的设计形式，但这根本就没有必要。进化自会对其造物精雕细琢。语法规则、词语、宗教实践以及其他许多人类文化中的基础特征也同样如此：无人设计它们，它们也不在“我们的基因中”，但即便如此，它们也经过了相当优秀的设计。

其次，为了这条人类独有的额外的信息高速公路、为了设计和传播的丰富媒质，我们付出了代价。和与我们一起生长的其他共生体一样，模因有其适应性，这种适应性在某种程度上独立于我们的适应性。忽视这一想法很常见，特别是当人们对宗教的进化做出解释时。“天哪！你在搞宗教进化论。你认为宗教带来了什么好处？它们一定有好处，因为很显然，每种人类文化中都有某种形式的宗教。”那好，每种人类文化中也都有感冒。它有什么好处呢？它是对自身有好处。我们不应该对发现那些不带来益处但却茁壮成长的文化自复制者而感到惊讶。这一见解取代了那种狭隘的思想，即文化革新跟基因革新一样，谁传播它们就能增强谁的适应性，由此平复了诸多文化进化理论间的竞争。模因是载有信息的共生体，就像栖居在我们身上的那些数以万亿计的互助共生体一样，我们不能没有它们，但这并不意味着它们都是我们的朋友。其中一些就是灾祸，我们最好没有它们。

第五部分小结

在这一部分，我试图表明，将达尔文的思想比作万能酸名符其实：它颠覆了整个传统世界，反对设计来自智能设计者的神来之笔这种自上而下的图景，代之以一个自下而上的图景。在这一图景下，笨拙、漫无目的的循环过程产生出大量的组合，该过程周而复始，直至这些组合具有了自我复制的能力，通过对其中最为优异者的反复利用使整个设计过程加速进化。

这些最早的生物体形成合力（一部主要的起重机：共生），从而导致多细胞生物（另一部主要的起重机）的出现，进而通过有性繁殖（又一部主要的起重机）使更为高效的探索成为可能，最终，它在一个物种当中产生了语言和文化进化（又是一部起重机），而这又为文学、科学和工程技术的传播发展提供了媒介，这些新的起重机崭露头角，反过来使我们以其他生命形式所不能实现的方式“追根寻源”，在方方面面反思我们是谁、来自何处，并借助戏剧、小说、理论模拟和计算机模拟以及我们引人瞩目的工具箱里越来越多的思考工具建模重演这些过程。

这一观点视野广阔、浑然一体，同时它又不吝提供各种精致、深刻的见解，可以说，它本身就是一个强有力的工具。那些还在莫名其妙地排斥达尔文思想的人得好好想想，如果他们只用老式的手动工具，便会发现自己的辛勤劳作远离了研究前沿，无法跟进对流行病、认识论、生物燃料、大脑结构、分子遗传学、音乐和道德这些纷繁复杂的重要现象的研究。

第六部分

关于意识的思考工具

INTUITION PUMPS

AND OTHER

TOOLS FOR THINKING

我们装备好几十样思考工具，终于迎来了“意识”这个许多人视之为全宇宙最让人困惑的现象。实际上，有不少人宣称，意识是“终极谜团”。他们说，我们绝不会理解意识；它将系统地证明，就算走到时间的尽头，科学和哲学付出的心血也终会落空。由于并没有充足的理由来相信这种智力上的障碍，我不得不说，这只是一厢情愿。还有一些人，他们不愿我们去揭示意识的运作，不想让我们的理解强加于

他们，他们劝我们早点放弃，因为那注定是徒劳。如果我们听取了他们的“忠告”，他们可就赢了，所以，不要理他们，让我们继续这场困难但并非不可能完成的探索吧。

53

对立的两幅意象
一幅错误的意象要用另一幅相反的意象来抵消

前面我已经介绍了数种思考工具，这些思考工具以这样或者那样的方式涉及心智的不同方面，如相信、思考等，但我一直把意识难题拖到了现在才讨论，原因在于，琢磨意识之际，人们总会夸大“意识是什么”这个问题，把自己弄得糊里糊涂。其实，我们本可以多花些精力来研究一下心智工作或运行的原理，把有关意识体验的那些永恒话题暂时放一放。我们扎下大本营不就是为了征服山巅吗？没错，不过若这么想，我们就已经想错了！心智中并不存在一座壮丽的意识之孤峰。这一传统观点至少可以追溯到17世纪的笛卡尔，然而与之相反，意识现象既非心智的“中心”也非“制高点”（Jackendoff, 1987; Dennett, 1991a）。

要抵消一幅错误的、充满诱惑的意象往往需要另一幅反意象，所以让我们从一个简单的想象调节器开始：回忆下科尔·波特（Cole Porter）的名曲《你是最高峰》（*You're the Top*），想想看，你怕不是山顶——你不是群山之巅，而是整片山峦；关于这片山峦，也就是你，你的所知所见并非是来自山巅的俯视，而是来自置身群山的多种视野。你差不多得把意识现象想成秃顶男人头顶边缘的那一圈头发。请牢记这一点。

接下来是另一幅反意象：意识不是像电视那样的媒介，可以对信息进行纪录或转换；大脑中并不存在这么一块把“一切信息都集中起来”的区域，像上演戏剧一样供某个中枢观察者品鉴，我将这个虚构的地方称作笛卡尔剧场（Dennett, 1991a）。与电视相比，意识更像名声：它是大脑中的名声或名流，通过这种方式，一些内容变得比其他内容更有影响力、更让人难忘。我不打算论证这一点（相关论证见Dennett, 1991a, 2005b），我只想为你提供一件简单的思考工具和一些友好的建议，用不用随你：但凡你把进入意识想作抵达总部，或想作无意识的神经尖峰信号转化为另一种东西，那么请你想想这些反意象，问问自己是否在错误地想象意识现象。

54

僵尸直觉
莱布尼兹的不合理推论

大多数人都有一种直觉，由硅、金属、塑料等材料制成的机器人绝不会以我们人类的方式拥有意识。意识必然涉及我们人类的生存、呼吸、身体以及大脑。这种直觉几乎不需要“泵”，因为它无处不在。有这种直觉的人也许是对的。不过，既然我们认识到，我们的身体和大脑由一些机器人组成，而这些机器人也是由机器人组成……一直到神经元层面以下，在那里，动力蛋白和其他纳米机器人不辞辛劳地忙碌着，维持着整个系统的运转，我们就会发现，这种直觉也许只是想象力贫乏的表现：人们总是在想一些简单的机器人，比我们人类的复杂程度差很多个数量级。

我的一位朋友试图一揽子打发掉我一开始的发难，他回应道，“我就是不能设想机器人有意识！”——“胡扯，你其实是在说你不想设想有意识的机器人，”我答道。你可能觉得在这上面较真有点傻，很不严肃。但三岁小孩都能设想有意识的机器人，比如，设想有意识的呜呜火车或有意识的圣诞树——所有心思敏感的小孩心中都会有棵孤单的小杉树，它渴望能有一个家。在观看《星球大战》的一个多小时里，几乎每个人都认为R2-D2和C-3PO是有意识的。我们从孩提时代起就已经这么做了，根本“无需细想”。只要我们发现有什么东西像人类那样活动，尤其是像人类那样交谈，我们自然而然就会将意识赋予它们，这是一种不可抗拒的想象。

说件有趣的事：20世纪50年代，神经科学家怀尔德·彭菲尔德（Wilder Penfield）在蒙特利尔做出了一系列开创性的工作，自那之后出现了许多例患者完全清醒状态下的开颅手术，手术中，患者能够在其大脑不同部位受到刺激的时候说出他们的感受。我想，任何一位试验参与者或观察者绝不会想说，“哦，天呐！这东西肯定不是人，他是僵尸。一定是僵尸，我们在里边看到的只有灰质。”我们当然不会这么想，你且

看且听，这些病人显然是有意识的。不仅如此，就算我们在某人与我们谈话之际打开他或她的颅骨，发现颅腔里塞满了芯片，这个人也显然是有意识的。或许这有点儿出人意料，不过我们马上会知道，有意识的机器人不但易于设想或想象，它们还真实存在。

一些哲学家认为，如果你采信意识的这种“纯行为”证据，直接跳到结论，你就被自己的想象欺骗了。他们的座右铭可能是“别信，别跳！”要证明一个人是否有意识没那么容易，也许他真是个“僵尸”呢——至少逻辑上有这种可能。这里所说的僵尸并不是我们在电影里看到或在万圣节装扮的那种巫毒僵尸。那些个行尸走肉很容易让人根据其行为和恐怖的外观将其同普通人区分开来。相比之下，哲学家设定的僵尸可以是我们可爱的同伴、派对中的明星，它们跟你平时接触的人一样充满爱心、阳光快乐、亲切自然。你最好的朋友就可能是僵尸。按照定义，这种僵尸在行为上与有意识的平常人类没有任何区别，它们只是有点儿“心不在焉”——没有内心生活和意识体验。从外部看，它们只是表现得像有意识一样。如果你同意这些哲学家，认为这是一个严肃的议题，此时面对逻辑上有可能存在的哲学家所说的僵尸，你怀疑如何才能为意识给出一个科学的、物质主义的理论，那你就中了“僵尸直觉”的圈套。(86)

不得不承认，和所有人一样，我也能很清晰地感受到这种僵尸直觉。我定神一想，意识确乎是某种东西，除了它为我们所做的一切、呈现给我们的一切之外，它还像是某种特别的个人独有的光辉，抑或是在机器人身上看不到的那种“我就在这里”的自觉——这种东西几乎不能看作“纯粹的”大脑物质活动。不过，我心里清楚，这种直觉并不可信。我觉得，它单单是一个错误，是想象力贫乏的产物，它根本就不是一种必然性的洞见。然而，要说服他人绝非易事，我们须得动用几个直觉泵来解开“僵尸直觉”的圈套。

在开始前，我们先将“僵尸直觉”在逻辑上的可能性同其他逻辑上的可能性做一番比较。逻辑上讲，你有可能生活在《黑客帝国》的世界里，你所看到的、你参与其中的一切只是一场虚拟现实秀，整个秀场设计得让你从容不迫，而你真实的身体正安静地呆在某种茧状的高科技容器里。逻辑上讲，碳原子可能不存在，科学家看到的碳原子其实是外星人驾驶的微型飞船，他们终其一生都在冒充碳原子。逻辑上讲，整个宇宙也可能是6千年前才诞生的，所有所谓的化石都在它们该在的地方，而光子就好像才从数光年之外的星系流入。逻辑上讲，我们的世界也可能在10分钟前才诞生，所有你有关过去的所谓记忆已经安装在了你的大

脑里。我们会在小说中读到很多这类有趣的逻辑可能性，但我们不会当真觉得我们的物理学、化学和生物学需要全面修改或丢弃。那么，有什么东西可以让“僵尸直觉”变得更为切实、更值得我们考虑呢？许多严肃的思想家认为存在这种东西。

所有设计用来产生“僵尸直觉”这类思考工具的直觉泵都有一个原型机，或许是几百年前的莱布尼兹发明了它。莱布尼兹与他同时代的思想家一样地聪明、富有创造力，他与牛顿各自发明了微积分，然而，他还是迷陷在自己发明的这件直觉泵上了。

> 假设有一台机器，它被造得能够思考、感觉，拥有感知，设想将这台机器的尺寸等比例地放大，放大到磨坊那么大，人可以走进去。然后，我们来检查它的内部：我们只发现其中环环相扣的零件，看不出感知到底怎样从中产生。可见，只有在简单实体（substance）而不是复合物或机器那里才找得到感知。（Leibniz, 1714, para. 17）

这个“可见”是哲学史上最刺眼的不当推理。对于这一结论，莱布尼兹并未给出任何中间论证，他认为这个结论太明显了，根本无需论证。回想下第22章中提到的威廉·贝特森，这位20世纪早期的遗传学家无论如何都想象不到基因是物质实体。贝特森无法想见，每个细胞里的一条双螺旋上会有30亿对碱基对，认为这简直是荒谬绝伦，同样，莱布尼兹也无法想见，一座“磨坊”里竟会有多达数万亿的活动部件。无疑，莱布尼兹会坚持，“单靠增加更多的零件”不会让你从机器变成拥有心灵的人，但这也只是他的直觉，无法给出证明。如果说达尔文、克里克和沃森揭示了贝特森的想象力之贫乏，那么可以说，图灵已然淘汰了莱布尼兹的直觉泵。只可惜图灵还未彻底做到这一点。至少现在它还尚未被淘汰。

我觉得，在恰当的时候，僵尸直觉会消散在历史当中，成为一件记录我们灵魂受困的稀奇古董。但我不觉得它会就此绝灭，它不再会保有眼下这种迷惑人类心智的形态，而是变得不那么有毒，尽管它会丧失权威，但仍然在心理上有强大的影响力。我们早就经历过这样的事情：地球就是现在看来也似乎静止不动，太阳和月亮绕着它转，但我们知道这些现象仅仅是表象，忽视它们才是明智之举。绝对静止的物体和惯性系中完全不加速的物体似乎是两码事，但我们已经学到，不要凭感觉下判断。我期望有一天，哲学家、科学家以及普通人都能笑对我们早先被意识现象迷惑的陈年旧事。“今天看来，为意识提供解释的种种机械论理论

似乎遗漏了什么东西，但我们知道那只是错觉。事实上，关于意识，所要解释的它们都已经解释过了。”

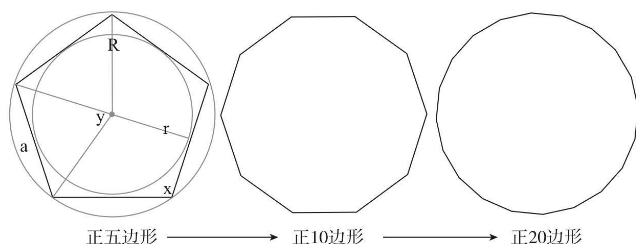
实际上，很多哲学家的思想实验仍在源源不断地为僵尸直觉提供支持，例如，塞尔著名的中文屋，这也是我创造“直觉泵”一词的灵感来源。不过，僵尸直觉马上就会在你们眼前被大卸八块。在此之前，我想稍稍细致地探究下“哲学僵尸”这个概念。

55

僵尸和殭尸

如何构想出一个哲学僵尸

每当有人说他可以设想哲学僵尸时，我们就有权问，你是怎么做到的？做出设想并不简单！你能设想超过三维的空间吗？你能设想空间曲率或量子纠缠吗？只是能想象还远远不够，正如笛卡尔教我们的那样，那根本不是设想。按照笛卡尔的观点，想象动用的是你的身体，它归根到底是一具机械，有着许多局限，比如看得不远、分辨率有限，角度、深度也有限；而设想动用的则是你的心智，它可是具有非凡洞察力的器官，不受机械规定的约束。为了说明两者的不同，笛卡尔还提供了一个令人信服的例子：正千边形。你能设想它吗？你能想象它吗？两者有什么不同？我们先来尝试一下想象。比如，从正五边形开始，想象正10边形。这有点儿难，不过你还知道该怎么做：将正五形的每条边从中点向外拉出一点，这样5条边就变成了10条边。要拉出多少呢？画一个外切于这个五边形的圆，从5条边的中点将它们拉至与这个圆内接的位置即可。重复这个步骤，你还能做出有20条边的正20边形。



继续重复7次，你就能得到一个正1 280边形。靠想象你几乎无法把它和圆区分开来，但论设想，正1 280边形显然不是圆，正如正方形不是圆一样，它当然也不是正千边形。想象一个内切于正千边形的圆，这个圆再内接一个正千边形，另一个圆再内切于这个正千边形……直到构成一幅类似靶的图案，告诉我，你能分清这幅脑海中的图像里，哪些是

圆，哪些是正千边形吗？根本不能，因为它们看起来都像是圆，同样的要求如果由设想来完成就没多大难度了。

笛卡尔并未要你干这些事情。在他看来，设想类似于想象，是某种直接的、偶发的心智活动：手到擒来，无需费心成像。你总可以在心智上把握到相关概念，如边、一千、规则的、多边形等，然后“嗖”的一下就全明白了。我始终怀疑笛卡尔对“设想”这种基本活动的想法是否恰当。如果你能这么设想，恭喜你，但我真的不行，正如在这词另外一个意义上那样，我也怀（conceive）不上宝宝。要确信我是否成功地设想出什么东西，我会反复摆弄相关想法，在头脑里彻底检查它们的含义，配以练习，直到最终熟练掌握相关工具。

在做这些精神体操的时候，我会格外倚重我的想象力，比方说，我会勘察脑袋里各种各样的图表和画面。总之，我利用笛卡尔贬低的“我的想象”实现了他颂扬的“我的设想”。你能设想弦理论吗？你是否发现，所有关于超弦、“膜”组成的多维空间的讲座易于理解，其逻辑一致性也易于测试？我无法理解它，正是因为我无法理解，我才不会说它匪夷所思、不合情理（Ross, 2013）。我并不信服这类讲座，但我也未自信到认为它们是胡说八道，也许是因为我的设想能力不够呢！至今我还没法设想弦理论的实情。我们不应在缺乏谨慎例证的情况下，过多地倚重可设想性或不可设想性给出的漫不经心的裁决。

贝特森曾说，物质性的基因“不可设想”，但他要是活到现在，就会轻而易举地学到如何来设想这种东西。要知道，小学生都了解双螺旋和它所有的碱基对横档，只要勤加练习就能轻松地设想这种现象。但是，再多的新信息和想象技巧都不会帮助我们设想圆的方，即一个正方形，它四条边上的所有点到其中心的距离都相等，也不会帮助我们设想最大的素数。

我相当肯定的是，哲学僵尸在概念上不融贯、不合情理，是个失败的想法。但你也别太信我的话，先来问问，要让你相信自己可以设想哲学僵尸，你需要做些什么？假设你试图想象你的朋友齐克“其实”是个僵尸。是什么让你相信，或更准确地说，是什么诱使你得出了这个结论？[87]齐克和僵尸齐克，他们的差别在哪儿呢？记住，齐克的所作所为不应让你确信他是或者不是僵尸。我发现很多人并没有想对，也就是说，在试着设想哲学僵尸的时候，他们不合时宜地遗忘或者漏掉了其定义中的某些部分。

所以，我们从僵尸中专门辟出一个亚种——我称之为殭尸

（Dennett， 1991a），也许这有助于让你看到自己的错误所在。所有僵尸当然都有非意识的控制系统，这种系统通过它们的僵尸眼球和耳朵从世界提取信息，并利用这些信息，避免它们走路撞上墙、在你招呼它们时会转过身来，诸如此类。换句话说，这种控制系统全都是意向系统。殭尸则比较特殊，它被赋予了一种可以监测自身内外活动的装置，因此具有一种内部的非意识的高阶信息状态，可以调控其他的内部状态。此外，自我监测装置还可以获取并使用自我监测状态的信息、对自我监测状态的自我监测信息，以此类推、直至无穷。换句话说，殭尸具备一种递归的自我表征能力，若能说得通，这种能力也是无意识的。正是凭借这一特殊的能力，殭尸才能参与到下面这类对话中：

你：齐克，你喜欢我吗？

齐克：当然，你是我最好的朋友！

你：你介意我这样问吗？

齐克：嗯，有点儿。这话有点儿伤人。我不喜欢你这样问。

你：为什么？

齐克：呃，我会觉得有点儿沮丧，或者说是被吓了一跳，也可能是从你那里听到这样的问题让我觉着吃惊。你为什么要这样问？

你：就让我这样问一下吧！

齐克：如果你执意要问就问吧。这整个对话都让我觉着别扭。

记住：哲学僵尸的行为举止与正常有意识的人没有区别，在其能力范围之内，它们也能开展类似上面这样的对话，为了表现出这类行为，它们需要递归的自我表征。它能够以其无意识的僵尸方式，对“它思考它何时想知道是否……的感觉意味着什么”进行“思考”。容易想见，如果当你以这种方式询问它的时候，齐克茫然不知所措，毫无反应，就会激起你强烈的怀疑，不过，这也让你发现，如果齐克是僵尸，它就不是殭尸。你应该始终确保：当你问及哲学僵尸是不是真的可能时，你是在思考殭尸，因为只有拥有递归自我表征能力的存在者才应付得了像上述对话这类日常交流，更不必说创作诗歌、构思新奇的科学假说以及表演戏

剧了，所有这些活动都在殭尸定义的能力范围之内。

只有当你想要细致地区分“正常”齐克与殭尸齐克，从而陷入麻烦时，你才真正开始设想哲学僵尸。否则，你就跟莱布尼兹一样浅尝辄止。现在，再向自己多问几个问题：你为什么在意齐克是不是殭尸？或者更个人化一点，你为什么会担心自己是不是或会不会变成殭尸？事实上，你永远得不到答案。

真的吗？齐克真的有信念吗？或者说，他仅有近似信念？——“别忘了，那种缺失了意识的信息状态在引导殭尸的生活，正如信念在引导我们其余人的生活。”就在此外，近似信念有了效力，用得着，是“真实的东西”，所以近似算子就用得不恰当。想象左撇子都是殭尸，比如我，而右撇子是有意识的，这可以帮助我们澄清这一点。

我：听说你已经证明了我们左撇子都是僵尸？我这辈子都想不到！我们可怜？你是怎么看出来的？

右撇子：哦，按照定义你们就是没有意识——这还不够糟糕吗？

我：槽糕？那要看对谁。话说无事不登三宝殿。你在干吗呢，想跟我一个殭尸对话？

右撇子：呃，对我而言，这里只是看起来好像有人。

我：对我而言也是这样的啊！毕竟，作为殭尸，我也具备各种高阶自我监测能力。我知道何时我在沮丧、何时我在痛苦、何时我很无聊、何时我很欢乐等等。

右撇子：不是，你只是表现得好像你知道这些事情，其实你什么都不知道，你只是近似知道这些事情。

我：我以为这是对近似算子的误用。你口中的我的近似知识跟你所谓的真正的知识根本就没有区别——你只是定义说，殭尸的知识不是真正的知识。

右撇子：但两者有区别，绝对有区别！

我：在我听来这完全是偏见。

如果这还不足以呈现和殭尸交朋友是个什么情况，那我就再试试其他的方式。说正经的，我们来考虑写部小说，讲讲一个殭尸被困在有意

识的人类世界，或一个有意识的人勇闯殭尸岛。你要设计什么样的细节才能让这个故事看起来合乎情理？或者干脆抄近道：读一部不错的小说，假设该小说的角色都是殭尸。那么什么情节才能暴露出角色的殭尸本质，又是什么情节才能否定该假设呢？小说家可以选择自己想要的叙述视角，例如，赫尔曼·梅尔维尔的《白鲸》选择了第一人称叙述，塞林格的《麦田里的守望者》也是如此：

> “叫我以实玛利。”
>
> “我要是为什么事心里真正烦起来，就不再胡闹。我心里一烦，甚至都得上厕所。只是我不肯动窝儿，我烦得甚至都不想动，我不愿随便动窝儿打断自己的烦恼。”

其他一些小说家会选择第三人称全知叙述。说来奇怪，第一人称叙述模式似乎更支持僵尸假说。别忘了，整部小说描写的只是殭尸以实玛利或殭尸霍尔顿的叙事行为（narrative behavior）。我们只看得到它们的外在表现，但它们究竟是不是殭尸只能从内心描写看出来！将第一人称叙述与第三人称叙述做个比较，例如，简·奥斯汀的《劝导》和陀思妥耶夫斯基的《罪与罚》：

> 伊丽莎白觉得慕斯格雷夫夫人及其家人都应该受邀共进晚餐，但她不能让晚宴显出一点异样——风格上不能有差异，仆人也不能少——她不能让这些从来都不如埃利奥特的人看出任何差池。虚荣心最终战胜了手头拮据的现实情况，伊丽莎白脸上又露出了悦色。
>
> 拉斯科尔尼科夫看着索尼娅，深深感到了她对自己的爱，奇怪的是，如此之深地被人爱着竟会让他觉得紧张、痛苦。真的，这种怪异的感觉简直糟透了。

看起来，作者让我们窥到了伊丽莎白和拉斯科尔尼科夫的内心世界，如此，他们怎么还可能是殭尸呢？但请记住：在人有意识流的地方，殭尸就会有无意识流。殭尸到底不该是奇迹，它们的行为受大量复杂信息的内部运作控制，受功能性类情绪的调节，如快乐悲伤、痛苦的类似物。所以，伊丽莎白和拉斯科尔尼科夫可能就是殭尸，奥斯汀和陀思妥耶夫斯基从常识心理学汲取我们喜闻乐见的语汇来刻画它们的内心活动，就像象棋程序员谈论电脑程序的迭代“搜索”和风险“判断”。殭尸

也会因社会地位不保而难堪、因爱情而神伤。

切莫忘记贝特森在想象力上的无能。每当我竭尽全力避开这个圈套，在我的背景假设里寻找破绽，留心我可能出错的地方时，我就总是会发现，意识这整个概念非常地混乱。例如，我可以想出两种，甚至7种或99种不同的所谓意识：左撇子有一种，右撇子有另一种，或许龙虾还有另外一种。但到目前为止，要想象这两种意识，我所能想到的唯一方法就是想象它们可以借功能上的差别加以区分：左撇子做不了X，而右撇子做不了Y，诸如此类。

然而，这些可区分的差别恰恰表明，我们讨论的根本不是哲学僵尸，因为按照定义，哲学僵尸和“真正有意识”的人之间没有任何外在可区分的差别。而且，到目前为止，若想为真正的意识明确地给出一个内在可区分的标志，那么不管怎样，这一标志就必须要求号称有意识的人动用“心智”，让我们以及她自己相信她有意识。但无论这种心智差别是什么，它都可能会在僵尸的“无意识流”那里有其对应的冒牌货。要是没有呢？——怎么会没有呢？所以说，我相当地确信，整个哲学僵尸的观念都是某种智识幻觉，是我们成长中的烦恼。试着克服它吧。我会在本部分靠后的地方为这项我们自己来说服自己的任务提供进一步的协助。

56

花椰菜的诅咒
感受质的意义之一

看你在狼吞虎咽地吃蒸花椰菜，我开始好奇你怎么会喜爱那种口味，只要尝上一小口我就有点儿想吐。也许是因为花椰菜的味道对你来说和对我来说是不一样的。对，一定是这样。自从发现即使同样的食物，在不同的时间品尝也会有不同的味道，我就更觉得上述假说很可信。比如吃早餐的时候，人们总会觉得第一小口橙汁是那么香甜，但如果中间来一点煎饼和枫糖浆，橙汁再入口时味道就平淡多了。这时候，抿一两口咖啡，橙汁的滋味就又恢复到了之前的甜美，是粗略地恢复还是精确地恢复？当然（叮！），我们就是要说或者思考这些事情，这当然（叮！）没什么问题，所以，……谈论一下丹尼特在t时刻品尝的橙汁味道如何当然（叮！）也不在话下。另外我们还可以再问一下，这味道是否与丹尼特在t时刻品尝的味道相同，是否与汤姆在同一时间t时刻品尝的味道有差异？我们把“事物对我们来讲是怎样的”称为它的感受质（qualia）。

这个“结论”看起来没有什么问题，但我们错就错在这里。在最后一步，可能是出于论证的目的，我们就把“感受质”从其他一同发生的各种状态中单独隔离了出来。人们假定，对于某人X来说，一杯果汁尝起来是什么样子可以从伴随物、附带原因或者附带产生的一些效果中区别出来。我们模模糊糊地想象，各种各样的事物在不同的时间对于不同的个体，看起来、听起来、感觉起来、尝起来、闻起来是什么样子，而不考虑这些个体具体是如何接受刺激，又在非感官层面受到了什么别的影响，也不考虑他们在感受之后做了些什么或倾向于相信什么。这种想象只保留了本质，而把其他所有的东西都一层层剥了下来，残留下了所谓的共有性质。我们的错误并不在于我们假设在实践中总是可以相当确定地进行这种提纯工作，根本错误在于，就没有这种最后残留下的性质，无论我们在提纯的时候多不确定。

在每一种感觉中，我们都在经受着各种各样的提取感受质的诱惑。我想象不出，或者说不知道，当然也不可能知道，对于格伦·古尔德（Glenn Gould）来说巴赫的音乐听起来会是怎样的。我已经记不起孩童时代我听巴赫的曲子时是什么样的感觉了。我也不知道当一只蝙蝠是一样怎样的体验（Nagel, 1974）。就颜色而言，每当抬头看见那块明净的“蓝”天，我也不知道你看到的那种蓝跟我看到的那种蓝是不是一样。日常生活中的各种案例会让我们不由相信，我们主观尝到的、看到的、闻到的、听到的各种特有性质都是实实在在的。然后，为了清楚地说明这些性质，我们就得通过哲学蒸馏对它们进行隔离提取。感受质就这样诞生了。

“感受质”是一个“技术性”用语，但对于我们来说它却再熟悉不过：“感受质”表示“事物对我们来说是怎样的”。一般来说，你最了解的是你获得的感受质。想象整个宇宙都是一团幻觉，是笛卡尔的恶魔虚构出来的，对你来说这种虚构就是你的幻觉体验的感受质。笛卡尔声称，要怀疑一切可以怀疑的事物，但他从来没有反思，他的意识体验里还有感受质，自己是通过这种性质才得知和了解到意识体验的。

看起来，感受质的定义已经说明得非常清楚了，哲学家们对这一主题也已经进行过相当多的分析和讨论，但严格地看，对于感受质的意思到底是什么，它到底意味着什么，我们还没有得出一致的结论。很多认知科学家大胆假定，认为哲学家们既然使用这个术语就一定知道它的含义，于是他们就把这个词吸收进了自己的工作语汇，这完全是一个策略上的错误。如今，关于感受质是什么或者不是什么大家仍然争执不休，但大多都是脱离经验主义层面上的一些讨论。几年前，我在一篇论文中（Dennett, 1988a）围绕着应该如何从根本上抓住感受质这个概念列出了下面四个要素，感受质是：

1．不言而喻的，

2．内禀的，

3．私人的，

4．我们可以直接接受的。

具体来说，感受质是这样的：（1）与内省相连，因而有些无法言说。“你得去了那儿才知道我是什么感受。”（2）不是关联式的，不是意向上的，也不是功能性的。比如红色常常能激发起人们的焦虑，但这

种主观倾向却不能算作红色的一种感受质。（3）“你得去了那儿才知道我是什么感受，但你去了也没用；这种感受属于我，并且只属于我！”（4）于你而言，你接受到的感受质要比任何其他性质都亲切熟悉。

按道理这还真是个不错的总结，只不过我当时写那篇论文的主要目的是为了论证：没有任何一个概念能够同时满足这四个要素。论文中，我还为这个概念提出了一些改进方案并附了充足的论证，但最后也没有给出一致结论。一般来说，一个普遍使用又备受关注的概念应该不会出现包含多种不同的含义且这多种意义还互不相容的情况，想一想生物学中的“基因”“物种”，还有科学中随处可见的“原因”，而“感受质”的处境却有过之而无不及。人们欢天喜地地迎接这份哲学馈赠的礼物，总希望它能在自己研究的学科中派上用场，没想到，它却是侵入到各门学科中的特洛伊木马。

在那篇论文中，不光有“花椰菜的诅咒”，我还展示了其他13种直觉泵，但我无意在此复述它们，因为这些年来，我又设计出了另外一些思考工具，它们更加有力，这才让我能在人们自以为准确地给出了“什么是感受质”而流露出自满后还得以有对抗之力。内德·布洛克（Ned Block, 1978, p. 281）借用路易斯·阿姆斯特朗（Louis Armstrong）的传奇名句“半开玩笑似地”避开了这个难以驾驭的问题。当被问到什么是爵士乐时，阿姆斯特朗回答：“你要这么问，你就永远没法知道。”这轻松好笑的回答淋漓尽致地表达出了人们的某种假定，它正中我下怀。如果我的对抗行之有效，那么你将看到，布洛克这个在任何社交场合都能蒙混过关的回答原来那么奇怪，根本毫无根据，就像一个活力论者在面对一位声称不相信“生命冲力”的人面前很打趣地表达自己对生命冲力的怀疑一样。

57

“美元活力”究竟值多少钱

感受质的意义之二

人们普遍认为，尽管我们可以造一个色觉机器人模型，把我们人类能体会到的所有视觉现象都展示出来，包括互补后像和色彩对比错觉等，尽管这个机器人利用与形成视觉现象相同的方法也能形成自己的视觉后像，但它也“只是个机器人”，它没有红色和蓝色的感受质。机器人的视觉功能只能将摄像机电子眼前有颜色的物体表现或标示出来，它没有我们人类视觉具有的那些额外功能。

托马斯·内格尔有一篇非常有名的论文，题目叫作《做只蝙蝠是一种怎样的体验？》（*What Is It Like to Be a Bat?*, 1974），文章为我们展示了比拟意识状态的一种标准方式，当然前提是真有意识状态这种实体存在。人们不难断言，“这种比拟明显不同于机器人对人类视觉后像的模仿！”但为什么这两种模拟会有那么明显的差异呢？也许是因为我们想象的那个机器人结构太过简单了，从它身上得出的有关计算机的结论不能涵盖所有计算机。按照我们的定义，如果感受质是体验的一种内禀性质，它应该独立于所有引发体验的原因及一切体验导致的后果，逻辑上，它也应该与所有意向性的性质无关，果真如此，感受质在逻辑上应该免于一切功能性分析。由此得出，再完美的工程设计也无法将感受质赋予机器人。然而，得出这个结论也终是徒劳一场，原因是根本就不存在感受质这种内禀性质。

看到这里，我打算将人类体验的感受质与货币的价值做个对比。很多美国人都天真地认为，美元不同于欧元或日元，美元有它的内禀价值。当卡通片里的观光者们问“这个究竟值多少钱”时，他们想知道的是这个东西价值几美元。再走得更远些：这些天真的美国人乐意拿美元兑换欧元，他们乐意利用美元的汇率兑换“贬低”其他货币，每次兑换都让他们觉得自己的美元真是与众不同。他们认为，每一块美元都有一些在逻辑上不受兑换原则约束的东西，而流通领域中的其他货币则没有。美

元有它独有的一些妙不可言的特点。当你思考它时，总会发现它戴着一个价值的光环，虽然没有以前那么耀眼，但依旧清晰可见：就让我们称它为“美元的活力”吧。按照书面意思，活力指美元非关系型、非意向性的内禀的经济价值。英镑、欧元等货币没有这样的内禀价值，它们只是美元象征性的替身，它们可以通过兑换美元获得相应的经济价值，但它们没有自身的活力！可怜的欧洲人！他们的货币没有内禀的经济价值！这叫人怎么受得了？他们哪儿来的动力去挣工资，里面只含有可怜的僵尸活力！我们挣的可都是美元，所以我们有幸能利用到它的活力。难怪美元会被那么多国家用作通用货币！原来连外国人也能感受到它的活力。

我们虚构出的那些美国观光客们就是这样认为的。从定义上看，经济学家们的理论永远也不会涉及美元的活力，因为没有哪种理论能解释所谓“内禀经济价值”。对于经济学来说，还有比这更糟糕的吗？钱币活力的存在让经济学出现了残缺，好在我们还没有找出太好的证据证实这种叫作内禀经济价值的东西真的存在。活力显然只是想象力的一种虚构，是单纯的美国人一厢情愿的产物，我要毫不留情地拆穿它。

在关于意识的讨论中，也有一些人跟那些想象出来的观光客们持有相似的观点。他们执意认为，自己直觉感受到的那些非凡的内禀属性才是所有意识研究必然的起点。这种自信应该被当作一种有趣的症状，值得所有意识学科去探讨，挖掘出一些数据，给出点什么说法。与此同时，经济学家和心理学家们也要着手研究一下，为什么“钱或美元有内禀价值”这样一种幻觉会有如此强劲的威力，能让那么多人都深陷其中。欧洲人也曾认为他们的钱，如法郎、马克、里拉等有内禀价值，习惯把外币换算成本国“真正的货币”，但后来启用了欧元，他们一度不知所措，陷入了尴尬。

我们现在应该对意识状态的一些特点做进一步的科学探究，如果思虑得当，它们还可能帮助我们更好地解释什么是意识，其实在此之前，生命的“奥秘”就是用这种方法才得以说明。活力论坚持，所有生命体中都有一些重要的神秘物质，它们是至关重要的活力。然而这只是想象力的一次失误，如今它已几近灭绝，虽然总会有一些怪人仍对它不离不弃。受生命“奥秘”成功解决鼓舞，我们可以继续对意识展开科学探索。但如果有一天，所有有关意识特性的问题都已清楚阐明，我们终于还清了智力上的债务，眼看着那些重大的问题就要慢慢消逝，到那时，那些曾经一度执迷的人就要开始喋喋不休，说他们早说过会这样，同时与活

力论者们一样，又要紧赶着消除掉之前受错觉误导遗留下来的某些论断。现在，问题来了，请问那些相信感受质是体验中内禀性质的人，你要怎样把自己的信仰与天真的美国人对美元价值的那些错误体验区分开来？换句话说，这些美国人想得对吗？美元真的有活力？任何人凭直觉就能知道？

58

卡普格拉先生的悲惨境况
感受质的意义之三

如果不是意识体验的内禀性质，那感受质到底会是什么呢？有天晚上，在喝了一瓶不错的香贝丹酒之后，哲学家威尔弗里德·塞拉斯对我说，“丹，感受质是那种支撑我们继续活下去的东西！”这个想法不错。接下来就让我们来细想一下，这样的感受质又是什么？为了看清楚问题，我会再启用一种直觉泵，它以近期的几项认知神经科学为背景，这些研究针对的是几种违背日常直觉的离奇病状：脸盲症（prosopagnosia）和卡普格拉妄想症（Gapgras delusion）。

脸盲症患者在大多数情况下都有正常的视觉，他们只是不能辨认人们的脸。他们可以区分出男女老少，可以分清楚谁是非洲人谁是亚洲人，但在面对相同性别与年龄的好友时，他们就迷糊了，除非听出了这几个人说话的声音，或是觉察到了其他一些可辨别的特征。如果面对的是一些照片，其中包括很多著名的政治家、明星，还有自己的家人和一些志愿参与的陌生人，这些脸盲症患者要想辨别出他们认识的面孔，只能靠撞大运。而非脸盲症患者的人则很难想象看着自己的母亲却认不出来的情景。很多人不相信我说的这些现象，非常肯定这都是我编造出来的。我谨希望你们能在那些让你觉得难以置信的地方更多地看到人类想象力的虚弱，而不是忙着去探究我描述得是否属实。更何况，医学家们早已开始研究这种困扰无数病患的病症——脸盲症。该词是由意为“脸”的希腊语单词prosopon与意为“不知道”的agnosia组合而成。

很多脸盲症患者会经历这样一件有趣的事情：虽然自己的意识不能辨认和识别别人的脸，但当面对熟悉或者不熟悉的脸时，他们却能表现出不同的反应，有时甚至会做出连自己都不知道的回应。或者，在面对询问时不能辨识，其实自己已经非常隐秘地分辨出来了。举个例子，我们为脸盲症患者出示一张照片和5个名字，让他们从中挑选出与照片相符的那个人名。虽然他们最终是随机选择的，但就在他们听到与图片相

匹配的那个名字时，他们的皮电反应确实就会出现一个很大的波动，这表示他们的情绪的确被唤起了。这一现象可以证明，那种隐秘的辨识力果真存在。

再来看下另一个简单的测试：要问玛丽莲·梦露、阿尔·戈尔、玛格丽特·撒切尔、迈克·泰森这些名字中，哪些是政治家的名字？你当然可以很快就做出解答。但如果这些名字配着错误的照片同时出现，你的反应就会明显减慢。这说明，尽管识别面孔与这个测试完全无关，你却在某种层面上是在识别那些人的面孔。

看起来，在我们的大脑中至少得有两种或两种以上独立的面部识别视觉系统：一种是脸盲症病人大脑中受损的意识系统，它在试验任务中无法为主体提供帮助；另一种是未受损的无意识系统，一旦发现面容与名字错配，它就会产生激烈的震荡。进一步还有测试表明：受损的意识系统位于视皮层，它的级别“更高一点”，而未受损的无意识系统位于边缘系统，相对来说层次就“低了一些”。如此总结脸盲症症状的多样性以及它们在大脑中所涉及的区域似乎太过简单了，但这样便于我们继续探索卡普格拉综合症这种更加奇特的症状。该病症由于法国精神病学家约瑟夫·卡普拉格（Joseph Capgras）在1923年首次阐述而得名。

患有卡普格拉妄想症的人会突然产生错觉，认为自己所爱的人，大多是配偶、恋人或者父母等，都是些被悄悄替换掉了的“冒牌货”。他们与一般的精神病人不同，从表面上看他们与正常人无异，但因为大脑损伤，他们的确会产生那些妄想，而且不管这些想法有多夸张、多不可思议，他们也都不以为然。有时，他们会重伤这些本是自己深爱的“冒牌货”，更有甚者还会杀掉他们。人们很难相信，一副受损的大脑怎么会有如此奇怪的反应？难不成撞一下脑袋就会把月亮认成是鲜乳酪？

认知神经科学家安德鲁·扬（Andrew Young）研究认为，卡普格拉妄想症的病理其实与脸盲症的病理正好“相反”。在卡普格拉妄想症患者的大脑中，大脑皮层上有意识的脸部识别系统总是完好无损，所以，在妄想症患者看来，眼前的人总是自己所爱的人的那副模样；而无意识的边缘系统却受损禁用，因而所有随情感共鸣产生的识别力也就此消失。鉴别力突然缺失掉了这些微妙的支撑可是件恼人的事，完好的系统在识别出熟人时做出肯定的地方统统都会被这个系统漏洞否决掉：妄想症患者真心认为自己看到的绝对是个冒牌货。他们从不谴责自己出错的感知系统，相反，他们抱怨这个世界，叫嚣着为什么事情会变得如此夸张、如此不可思议，他们从不怀疑“是不是自己头脑中非意识面部识别系统的

识别力出现了问题”。这部分系统对我们所有人都有着巨大的影响力，而一旦它的认知需求出现了饥渴，妄想症便随即发作，它会屏蔽掉其他所有系统的反应信息。

这一假设由海登·埃利斯（Haydn　Ellis）和安德鲁·扬在1990年首次提出，之后，扬、克里斯·弗里思（Chris　Frith）及其他一些神经科学家还对它进行了详细的论证及阐述，在这里，我不再多加赘述。我想要做的是，利用这种能够拓展想象力的特殊认知神经学为我们的心灵开辟出一种不可即但可望的可能性。下面是我虚构出来的一个有关卡普格拉先生的案例，起这个名字是为了提醒大家随时注意卡普格拉妄想症的那些真实表现。为了探索感受质对意识的影响，案例中引入了大量哲学类的直觉泵，虚构了一些意识上的病症。

卡普格拉先生靠着给心理学和精神物理学实验当被试，过着说得过去的生活，所以他对自己的主观心智状态相当了解。一天早晨，刚刚睡醒的卡普格拉先生一睁眼便绝望地大叫了起来：“哦！出问题了吧！为什么这个世界变得……这么奇怪，这么可怕，一定是有什么搞错了！我还要继续活下去吗？”他使劲揉了揉双眼，再次小心地张开眼来看，世界还是那副扭曲的模样，有几分熟识却又难以描述地陌生。有人问他：“抬头向上看，你会看到什么？”“蓝天，天上还飘着几朵绒毛状的白云，初春的枝头上罩着一层鹅黄的嫩芽，一只鲜艳的红雀栖息在枝头，”卡普格拉回答。看来他的色觉一切正常。就进一步检测结果来看，他通过了标准的石原式视力测试，他没有色盲症状，色视觉正常，而且他还正确分辨出了几十条孟塞尔色卡。所有人都断定，单在色觉上，我们可怜的卡普格拉先生没有任何问题，只有一位研究者克罗马菲尔博士，执意坚持要再测试一遍。

克罗马菲尔在人类色彩偏好、颜色引发的情绪反应和不同颜色对人的注意力、专注力、血压、脉搏、代谢活动及器官活动产生的微妙影响等方面都有很深入的研究。在过去的6个月时间里，他已经积累了许多卡普格拉先生在所有测试中的反应数据，包括正常的和非正常的，想看看两组数据间是否会存在些许变化。果不其然，在经过了重新一轮测试之后，克罗马菲尔博士发现了一件很有意思的事情：之前卡普格拉先生在面对蓝色时产生的情绪波动和器官反应现在换成了面对黄色时产生的反应，同理，之前面对黄色时才产生的情绪波动和器官反应，现在换成了蓝色。他以前喜欢红色、讨厌绿色，现在则恰恰相反，同样发生逆转的还有他对其他颜色的喜好。美食的颜色开始让他作呕，所以他不得不

躲到黑暗处吃东西。以前喜欢的颜色搭配现在被拉入了黑名单，而与之对立的那些色彩组合却突然变成了他的最爱，诸如此类。以前，鲜艳的粉红色会让他脉搏加速，但如今，虽然他仍能认出这种颜色，但对他起到的却是镇定作用，他很惊讶别人竟然认为这种颜色“鲜艳”。相反，这种粉色的互补色石灰绿，之前起到的是镇定作用，现在却轻易地就能挑逗起他的激情来。

你发现，在卡普格拉赏画的时候，他眼睛飞快扫视画面的轨迹与以前也大不相同，原因是画布上的颜色对他注意力的抓取和凝视偏向也产生了微妙的影响。以前，身处在一个亮蓝色的房间会抑制住他的心算能力，而现在，这种抑制他心算能力的颜色却变成了亮黄色。

简而言之，尽管卡普格拉对色觉测试中的诸多问题都毫无怨言，他准确地说出了颜色的名称，按规定通过了所有颜色的识别测试，对，他完成得相当出色！但在这个过程中，在两次测验中，他对色彩的情绪及注意力的反应却发生了180度的大逆转。卡普格拉身上到底发生了什么？面对同事们的惊奇和疑问，克罗马菲尔博士简单解释道：他的色彩感受质一下子被完全倒转了，只有较高层次上辨识颜色的能力没有受到影响，所以他还能区别出各种颜色的名称，这正是感色机器人能具备的那种能力。

接下来我们再往哪个方面研究？卡普格拉的感受质真的倒置了吗？案例是我们想出来的，貌似可以用任意我们喜欢的方式应对它。多年来，让哲学家们津津乐道的这种虚构出来的案例还有很多，其中每一个他们都认真对待，因为在它们之上，还有一些更深刻的理论问题。所以我们绝不能草率对待这个案例。首先要问的是，该案例合理吗？也许这要取决于我们讨论的是哪种合理性。它在逻辑上可能吗？在生理学上是否也可能？这是两个完全不同的问题。后一个问题看似离我们的哲学关注太远，就常被哲学家们无情地忽略，而在这个案例中，他们慢慢转变了看法。

我发现，其实我们根本无法否认这个案例在逻辑上是可能的。根据描述，卡普格拉一部分能力失常，剩余的能力保持不变，本该交融在一起处理的能力在他那里变得彼此孤离，然而，与脸盲症和卡普格拉妄想症患者比起来，这一点在卡普格拉身上体现得是否过于明显？我不确定，对卡普格拉在生理学上的设定是否像下列真实案例一样说得通。

在一些研究充分的案例中，有人可以清楚地分辨颜色但说不出颜色的名字，即他们有色彩命名障碍；有人色盲却不自知，能很坦然地跟大

家谈论各种颜色，却不知道所有颜色的名称都只是他们自己随机猜测出来的。卡普格拉能够辨认并正确说出各种颜色的名称，在这一点上他与卡普格拉妄想症患者一样，但他对色彩的微妙体会却全部变形。而成就人类个体倾向的也正是这些微妙体会，它们决定着我们会欣赏哪些图画，会描绘哪个房间，会挑选哪种颜色组合。在卡普格拉身上发生改变的正是左右我们生命意义的那些色彩效应，按照塞拉斯的说法，是他的色彩感受质。

如果我们真的要问卡普格拉是否知道自己的色彩感受质已经倒置，可能的回答有三：是的，不是，我不知道。他的答案会是哪一个？哲学家们已经公布并非常详细地讨论了一些有关感受质倒置的例子，而将我设置的案例与其相比，最恼人的创新之处在于，卡普格拉可能不只是遭受了感受质的倒置，而且他对此事始终一无所知。请注意，克罗马菲尔博士是针对同事们的怀疑提出这个假设的，而卡普格拉同样希望缓解自己的疑惑。但是，他不仅没有质疑自己的色彩感受质是否存在问题，相反，他对自己的色觉基本满意，因为标准色觉测试进行得相当顺利，研究人员对这个结果也同样满意。这样的设计会稍微有些让人不安，因为在哲学著作中，类似这样的行为型自我测试一般都是不切题的：这些测试肯定（叮！）与感受质无关，更有甚者，这些测试的特点就是不能对感受质问题有所匡正或者阐释。但是，我对这一案例的变异却能避免哲学家们设定上的一个缺陷：他们忽略了人们心灵上的一种倾向——想在那些测试中不断确定，自己的感受质没有发生变化。

当“情感上”发生变化时，我们的感受质是否保持不变？如何回答有关感受质的这些确切的问题，哲学家们的意见出现了分歧。考虑一下味精这种提味儿的调料，它无疑能让事物尝起来味道更加鲜美，但究其原因，是因为它改变了食物给予我们的感受质，抑或只是，它提高了我们对已品尝到的感受质的敏感度？这里亟需澄清一下感受质这个概念，它关乎的既不是“味精作用在舌头的哪个部位”这种经验主义的问题，也与被试自我报告中出现的那些“人体对味精的各异反应”不同，如果不把这个概念搞清楚，那么我们在基础的神经层面和他者现象学等诸多方面的发现就会整个儿变得模糊起来。我只是很想知道，哲学家们是在什么意义上使用“感受质”这个词的，他们能够分辨出主体伴随感受质的改变而在反应上产生的所有变化吗，或者，反应系统中也许有某种特殊的子结构可以有效锚定感受质？企图去改变一个人对于某种特定性质的审美观点或反应，这种想法会不会有些荒诞？“感受质”的定义不只是有点儿暧昧，它简直模糊得一塌糊涂，总是在两种甚至多种根本不同的想法间捉

摸不定，所以，解决好有关它定义的上述类似问题迫在眉睫。

卡普格拉的色彩感受质果真倒置了吗？对于这点，很多哲学家认为我给出的细节描述还不够充分。在案例中，我描述的基本都是他的行为能力：他会识别，会区分，会正确地说出颜色的名字，也会在其他一些地方犯“错误”，而对他的主观状态却避而未谈。当他在看一颗熟透的柠檬时，我没有说明他是否体验到了一种内在的主观上的黄色，或者说是内在的主观上的蓝色。可问题在于：我并不认为“内在主观上的黄色”或者“内在主观上的蓝色”就说出了卡普格拉体验的真实性质。如果卡普格拉在回答提问时加上了这句：“我依旧会认为成熟的柠檬是黄色的，可见，内在主观上的黄色应该囊括在我的体验之中。”这能解决什么问题吗？我们能肯定他真的清楚自己在说什么吗？我们应该相信他吗，或许他只是陷入到了一个不值得他如此深信的哲学理论里？

在这些案例中，哲学模式的标准化使用有这样一个主要缺陷：哲学家们倾向于把正常人面对色彩及各种事物时表现出的那些能力和性情看成是一整块大砌砖，不可能被分解或者分离成一些彼此独立的亚能力或亚性情，这让他们很巧妙地绕开了下面这个问题：感受质是否得安放在某个子集或者专门的设置中。举个例子，哲学家乔治·格雷厄姆（George Graham）和特里·霍根（Terry Horgan）认为，“对现象的直接认识使现象生出了自身的特征，认识是一个人识别或辨识能力的经验基础。”他们是怎么知道这种“直接认识”就是识别或者辨识能力的“基础”的？这么说来，脸盲症患者能够直接认识自己眼前的这些面孔，或者至少他们能直接认识眼前这些熟悉面孔的“视觉感受质”，但他们不把这些直接认识识别为他们面对亲朋好友的面孔时所体验到的感受质。再次回到塞拉斯的观点，感受质是那种支撑我们继续活下去的东西，那么它就可能不是我们日常认识颜色的“经验基础”，也不是我们区别颜色、指名颜色的“经验基础”。

59

听声辨牌
感受质的意义之四

哲学家大卫·查尔莫斯（David Chalmers）在他1995年的一篇著名论文中将意识问题分成了“简单”问题和困难问题两类。在我看来，即便他的那些“简单”问题也相当有难度，下面让我们来看一下其中几个有关意识的极具挑战性的问题：

1．意识如何能使我们谈论自己的所见所闻？具体说来，从大脑感知系统得来的信息是如何在大脑语言区域的控制下形成记录或者提炼出答案的？

2．在例行活动过程中，这些活动几近“在睡梦中都能做”，为什么我们的意识能在遇到麻烦时自动启动，同时，它又是怎样帮助我们解决掉这些问题的？

3．我们可以同时有意识地追踪几个独立运动的物体？我们是怎么做到的？（答案是，至少四个。）

4．已经“话到嘴边”——你知道你很清楚那个东西，几乎就能说出答案却说不出来，这是怎么回事儿？

5．为什么你能意识到那是一个笑话并觉得它好笑？（请参见Hurley, Dennett, and Adams, 2011。整本书都是这个问题的答案。）

对于查尔莫斯来说，这都只是些简单的问题，因为在大脑中进行信息处理、执行注意力指向功能，在清醒时开启追踪、提醒和回忆等等，它们牵扯的只是意识的认知功能。尽管它们之中也有些问题的解答原理非常繁难，但毕竟还能通过实验去检测和改善，而且，在解决这些“简单”问题的征程上，我们早已取得了不小的成就。例如，我们已经成功构造出一些相对比较简单的计算机模型，它们能把与之对应的大脑意识活

动精准地复制出来。由此可以断定，大脑在思考这类简单问题的时候没有借用奇异能力，也没有使用什么超自然的力量。在可预见的未来，这类现象都能由计算机一一展示出来。

对于查尔莫斯来说，与“体验”相关的问题才是那些困难问题，拥有意识是什么体验，这难以形容又不可分析。机器人的行为看似是有意识的，它们可以回答我们提出的问题，追视所有能见的动点，有“欲言难吐”的时候，也能重新组织出语言，它们可以在该笑的时候笑，在需要的时候不自觉地变得一脸茫然或者哑然失声。然而事实上，这都只是些无心之作。机器人就像前面提到的僵尸，它们没有你或者我这类正常有意识的人所享有的那种内心生活。

按照查尔莫斯的意思，亲爱的读者，你和我，我们从一睁眼起床开始就清醒地知道我们有意识。一个哲学僵尸不会明白这些事情，因为它没有所谓的清醒，也没有内心生活，它们只是从表象上看是有意识的。它们非常虔诚地坚持认为自己与你我一样都是有意识的，即使用测谎实验来检测，也看不出有什么差池。可惜它们终究是些僵尸，它们根本没有意识！神经系统科学家通过用功能磁共振成像（fMRI）这类机器对它们的大脑状态进行检测，就能很清楚地认识到这一点。看来，要清楚地将一个有意识的人和一个僵尸分开，这的确是一个困难的问题，如果它是个问题的话。而在这个问题基础之上，如果继续追问“这种不同是怎么产生的？”就形成了一个更加困难的问题，这才是查尔莫斯的“困难问题”。包括我在内的一些人认为困难问题只是查尔莫斯自己用想象虚构出来的一个东西，而多得惊人的另一些人则坚信，有意识的人与一个完美僵尸之间的确有区别，而且这种区别非常重要。

回顾这一奇特的景观，一些人认为，查尔莫斯所谓的困难问题不一定存在，而另一些人认为，否认下面这一点甚是可笑：对于一个有意识的存在体来说，没有什么比他自己的意识更明显、更直接，那是他们独有的，可直接觉察到的。我们仅凭意识的这个特点就有理由蔑视到目前为止所有的科学认知，所以它真真儿的确就是那个最困难的问题。这两种立场非此即彼，无法调和。多年来，我一直竭力摒弃这种直觉知识，不管它们多有说服力，我敢肯定，“存在着困难问题”这一观点绝对是个错误，只是我目前还给不出明确证明，或者，即便我给出了证明，也未必能引起重视，因为持这一观点的哲学家们早已准备好了他们无懈可击的“理论根基”——这是人类的直觉，这种直觉显而易见、不可否定，没有什么论证可以撼动它，更不会有什么东西能把它击碎。所以，我不会

用理性论证去证明这个错误，它远远超出了理性的范围，我不能犯这种战术上的错误。

这总让我想到那些刚看完精彩的舞台魔术表演的人常有的一些由衷的旦旦之言。魔术师们都知道，人们在回想一些伟大魔术的时候往往都会夸大其词，那些有震撼力的、让人迷惑的瞬间总能让人记忆犹新，人们总是执意认为自己没有遭受魔术师的愚弄，自己真的看到了一些神奇的景象。有的人是那么相信魔法。而回顾第22章《神奇组织》，李·西格尔在谈到“真正的魔法”时却说，“真正的魔法不是真实存在的魔法，那些可现实操作的、真实的魔法也不是真正的魔法。”

对不少人来说，意识就是“真正的魔法”。如果你谈论的东西还不够让人觉奇思妙想难以置信，那它就一定不是意识，它的神秘超出所有人的理解。科学记者罗伯特·赖特（Robert Wright）也很简洁地表达过类似的观点：

> 显然，这里的问题是，为什么有人会宣称意识“同一于”物理上的大脑状态？而丹尼特等人越是不停地向我解释这句话的意思，我就越是确信，“不存在所谓的意识”才是他们真正想要说的。（Wright, 2000, p. 398）

大脑里没有什么装着意识的锦囊，就算有，它装的也未必是真正的意识。有些人深知这一点，但还是不禁要夸大那些意识现象。这也是为什么我会在我的书《意识的解释》里花那么大的篇幅为意识消脂，让意识，真正的意识，恢复到原本紧实的状态，从而告诉人们，意识现象其实远没有他们想象的那样壮观。但是，这种消脂运动着实刺激到了很多读者，他们开玩笑地说，我的这本书应该取名为《被解释掉的意识》，赖特也提议，该把书名改成《被否定的意识》。对于那些已经陷入膨胀的意识观念还不自知的人，我只想给出个侧击，让他们能注意到我从纸牌魔术界引入的这个美妙又有些扰人的类比——听声辨牌，借此驱逐掉他们的自满。

1938年，约翰·诺森·希利亚德（John Northern Hilliard）在《纸牌魔术》（*Card Magic*）中写道：

> 拉尔夫·赫尔（Ralph Hull）先生是来自俄亥俄州克鲁克斯维的一位著名纸牌鬼才，多年来，他不仅用一系列纸牌魔术迷惑了普通公众，也让很多专业魔术师、纸牌专家和业余魔术师

们都陷入了困惑，他得意地将这套魔术称作“听声辨牌”……

赫尔的魔术大致是这样的：

> 他说：“小伙子们，现在，我要给你们演示一个新的招术，它叫‘听声辨牌’。这副牌可以很神奇地发出乐音。”赫尔将一副纸牌凑到自己耳边快速地翻动，自己静静地听着纸牌发出的嗡嗡声。过了一会儿，他说：“伴随着这些乐音的微微震动，我就能听到并感觉到每张牌所在的位置。我能挑出你想要的任何一张纸牌。随后赫尔把这副牌抹成一个扇面出示给大家，一位观众从中取出了一张，记下了它的牌面，然后又将这张牌放回了它原来的位置。“现在我就来听听这副牌，它都跟我说什么了？有提示我的声音？”嗡——，嗡——，赫尔在耳边弹洗了几次牌，然后又例行纸牌各式各样的规定操作，一派炫技之后，他果真亮出了之前观众选定的那张牌。

赫尔曾特意演示给同是魔术师的朋友们看，一遍又一遍地重复，看他们能否瞧出其中的端倪。可惜，没人做得到。其实，重复演示给观众看是纸牌魔术的大忌，当然，像这种伟大的招术则无须有什么担忧。很多魔术师想要买下这个魔术，可赫尔不卖。赫尔晚年将这个魔术的解密交待给了自己的一位好友，希利亚德，希利亚德后来将它写入了自己私下印制的书里。关于自己的魔术，赫尔曾说过：

> 多年来，我一直在表演这个魔术，数以百计次地将它展示给所有最好的专业魔术师和业余爱好者们看，但没人能看出其中的蹊跷……小伙子们在找寻答案时都太用力了。

很多伟大的魔术都有这样的特点：在你还没有意识到它是否开始时，招术就已经过去了。听声辨牌这个招数就全在它的名字“听声辨牌”（The Tuned Deck）里了，更确切地说，就在“The”这个单词上！就在他为热切的观众们演示自己的新招数，就在公布新技法名称的一刹那，招术就结束了。

赫尔用这个简单的方法吊足了观众们的胃口，中间加上一些明显的小失误，穿插着误导性的台词，还伴随着“嗡——嗡——嗡”纸牌发出的声音，然后，他使出一个相对简单又被大家熟识的纸牌招术A（现在，我要揭开以往总是挡在前面的那层帷幕了，你会看到，那些专业操作上

的细节反而不那么重要）。赫尔的观众都是一些很懂行的魔术师，他们能觉察出舞台上要演示的可能就是A招术，于是他们开始变得顽劣、不配合，极力阻碍这一招术的进展，他们想用这种方法来验证自己的猜想。观众席上已经出现了一些抵抗力量，而就在这个时候，赫尔转而又将这个招术“重演”，只不过这次他要展示B招术。观众席上又是一阵混乱，一番比较之后，大家开始推断，他这次展示的是不是招术B？观众们又开始想尽方法阻碍赫尔的表演，因为他们想再次印证自己的猜想。赫尔继续演示“这个”魔术，只不过这下又换成了招术C，当人们再去印证他采用的是不是C招术时，赫尔继而又转向了招术D。当然，此时他也可以再回到招术A和招术B，因为观众已经否决了这两种猜测。他继续这么切换下去，重复多次，每次在大家验证猜想之前切换到下一个即可。

赫尔知道，他的每一个招术观众们都非常熟悉，他只是比别人多明白一件事：只要加上一个定冠词“The”，就能让这堆常见的纸牌招术顿时销声匿迹，这就是“听声辨牌”。正如赫尔对希利亚德所说，

> 他走的是这样一个套路，让观众脑中的想法接二连三地破灭，最终，他们就会放弃再尝试去解开这个谜题。

赫尔的听声辨牌引入的是一个非常普通的词“The”——老天保佑！这个温和的单音节词迷惑住了所有的专业级观众，消磨了他们的意志，让他们无法实现“跳出系统”。这个“The”总让观众们觉得自己就要发现那个大阴谋了，但又不断出乎他们意料。实际上，大阴谋可不止一个，而是一个接一个地出现，直到观众深深地陷入了泥潭，无法自拔。

我认为，查尔莫斯在向世界提出他的“困难问题”时，同样也使用了这套手法，当然他并非有意为之。“因难问题”真的存在吗？或许，困难问题和“听声辨牌”差不多，也是一大包招术，它们是查尔莫斯定义的有关意识的那些简单问题？它们只需要一些很平凡的解释，不用在物理学上有任何革新，也给不了我们什么意外惊喜。只要用认知科学的标准方法去研究，历经艰辛努力之后，它们终将解决。

我证明不了这世上真的没有那些“困难问题”，同样，查尔莫斯也无法确定一定就有。查尔莫斯只是直觉认为有这么一类“困难问题”，如果这个直觉能引出新的惊人启示，或者能为更多的疑惑提供解释，那我们还有必要尝试和他一起围绕这个直觉去建构一套新的意识理论，但它实在缺乏自身的理论活力，尽管要否认它也很难。

大家知道的意识的效用已有很多，而且会越来越多，它们有的普通，也有的比较独特。统统把它们都记录下来实为不易，而事实上，我们要格外小心别犯算术错误，自以为在整合好所有的简单问题之后，还剩下些不明来历的残余。这些残余可能早已在一些平凡的现象中安顿了下来，只是我们没注意罢了。一般来说，我们对这些平凡的现象都能给出清楚的解释，即使没有得出最终解释，至少探索的道路清晰可见，那为什么我们还会犯算术错误而不自知？我们会不会把有些现象数了两遍，或者已经解释了某现象，后来又忘了，从而没有把它们从这张“有待解释的现象”清单中抹去。思考一下，你面对卡普格拉先生，从某些地方看出了他的不对劲，但酿成这些差错的是两条完全不同的路径：

A．他的色彩感受质的美学反应和情绪反应已经完全倒置，但他的感受质本身保持不变。

B．他的色彩感受质是倒置的，但保留了对颜色的辨别能力，也能说出各种颜色的名称。

就像那些疯魔的观众揣测赫尔的招术那样，你可能会辩解：“A不对，要说他的色彩感受质没有发生变化，我们唯一能给出的理由就是，他的命名和辨识行为都正常，而这两者根本不可能对感受质有所证明，因为它们都是认知与功能上的事实，本与感受质彼此独立。B也不对，因为他只是对色彩的反应发生了变化，卡普格拉并没有抱怨相同的颜色为什么看起来先后不一，而是相同的主观色彩为什么不再对自己产生同样的吸引力。或许卡普格拉的色彩感受质真的已经改变，也或许没有，但请你记住，没有什么实证的方法能帮助我们做出判断！这真是一个困难问题！”

而这些辩解却忽视了这样一种可能性：感受质在A和B中不发挥作用。我们可以看到，鉴别装置在A和B中均正常发挥作用，但对于它生产出的信息，卡普格拉做出的反应却是倒置的。感受质充当的是一种无可抵挡的居间力量，它为情绪反应提供必要的基础、原材料和平台，所以倒置可能在两个地方发生：一是在感受质“呈现”给鉴别装置之前，一种是在感受质“呈现出”之后，鉴别装置要对已呈现的感受质作出回应时。这可真是一次承载颇多的“呈现”呀！我们知道，比如说，在知觉过程中，消极反应，例如惊恐、害怕等会被较早触发，从而“渲染”了后续的知觉输入，由此我们得出，可能是情绪反应让感受质带上了卡普格拉那样的主观特征，而不是感受质的“内在”本质引发了情绪反应。但是，如

果问题真的出在对知觉输入的情绪、审美、或者情感反应这个环节，那么在感受质方面就没什么文章可做了，殭尸就会和有意识的人一样因为知觉反应的倒置而灰心丧气。

听声辨牌这个故事为感受质其他的那些直觉泵带来了哪些提示呢？事实上，这个真实的案例告诉我们，即便博学睿智的专家也会仅仅因为问题呈现的方式而捏造出一些虚幻的问题。这样的事情时有发生，也将会继续发生。由此引来一个新的问题，我们有责任提供证据：怎么才能看出我们是否已陷入了一个类似“听声辨牌”那样的骗局？我并不是为感受质的问题盖棺定论，而只想建议那些深信僵尸直觉的人再重新考虑一下这个概念，看看它是不是果真那么“显而易见”。

60

“中文屋”错在哪里
理解可以从一堆能力中冒出来

20世纪70年代末，人工智能喧嚣一时，夸大了该领域的现状和前景：思考机器就在眼前！而约翰·塞尔却坚信自己看透了其中的破绽，还构筑了一个思想实验予以说明。1980年，在文章《思想、大脑和程序》（*Minds, Brains and Programs*）中，塞尔提出了“中文屋”这个著名的思想实验，试图证明“强人工智能”不可能出现。他这样界定“强人工智能”：“给计算机适当的编程让它拥有认知状态，只要运行程序就可以解释人类认知。”（Searle, 1980, p. 417）随后，又进一步阐明：“适当编程的计算机，如果输入输出均正确运行，那么它就能拥有完整意义的人类思想。”（Searle, 1988, p.136）

众所周知，1980年，在认知科学核心期刊《行为与大脑科学》（*Behavioral and Brain Sciences*）和它的网站论坛上发表的文章都有这样一种特殊形式：围绕一个问题会有多篇长篇的“专题文章”，之后还附有数条领域内专家们的评论及作者们的回应。网站论坛极力追求的是跨学科讨论，各个学科的代表专家纷纷发言，这就为读者们提供了及时且有用的反馈。一篇专题文章能否引起其他专家重视，他们是如何认真对待这篇文章的，这些可以帮助你预测一篇文章能否可以为己所用以及如何利用。另外，与自信的讨论者们发生激烈对峙，无视彼此，各说各话，或是参加到一个拥有最强实力的学术战斗团体中，这些体验也能让你了解到很多跨学科讨论上的困境。塞尔的那篇专题文章引来了一系列激烈的反驳，包括我自己在内（Dennett, 1980）。也正是那个时候我发明了“直觉泵”这个术语，协助我揭露出塞尔那个思想实验给大家带来的误导。

虽然“直觉泵”是我发明的术语，但也借助了侯世达的力量，我是在与他一起讨论塞尔文章的时候才萌生了有关“直觉泵”的想法，后来我们还将塞尔的文章连同评论一起放入了我们的文集《心我论》（*Hofstadter & Dennett*, 1981）。我对塞尔的这个思想实验感兴趣，出于两个原因：

一方面，这个论证确实有误导性，它在欺骗读者；而另一方面，它又是那么有说服力，读起来实在让人感到愉悦。那么，这个思想实验是怎样进行的，为什么会如此运行？我们曾用逆向工程的方法探索过这些问题，候世达建议我们要“转动所有的旋钮”，去看看这个思想实验是如何运作的。这个故事可以经受住较大幅度的“变形”吗？还是它得严格取决于本应是随意设置的细节才能成立？[(88)]

我必须承认，在30年前，塞尔中文屋的思想实验确实有着非凡的感染力和影响力。它堪称经典，在成千上万的本科教材中出现过，至今被讨论过无数次。我在自己的课堂中也曾多次引用它，我想研究它是如何运作的，还想试着向人们展示出它的问题所在。[(89)]

为此，我要先展示这样一个直觉泵——吊杆托架，它存在一些问题，所以要小心使用，以免丧失想象力。对于我即将展开的这项细致的工作，我想你们听过后就会默默把眼睛瞥向别处，甚至向我“哼”一声，实在不想看我毁了“中文屋”这个设置，你们多么喜欢这个结论——吁，根本不可能有强人工智能存在！所以此时，你们目光呆滞地望着远方，盼着谁能再悉心研究出一个生动有趣的论证。我看不起你们，你们贪恋的只是好不容易有个著名的哲学家写了一篇文章，为你们的想法提供权威论证，为此你们高兴极了。你们真正关心的不是细节，只是结论，这是一种逃避，一种反理智的逃避行为！

但随后我也发现自己做了同样的事情，我必须重新考虑我的判定。我承认，我一直都觉得量子力学莫名让人讨厌，因为再怎么努力我都无法把它弄明白，有时候，我宁愿它是无中生有！我深知它对一些现象的预测和解释让人心服口服，其中包括很多日常现象。例如，光的反射和折射、我们用眼看事物时视网膜上蛋白质的运作等等。量子力学是一门核心科学，但即便是量子力学的专家们也未必能将它完全弄懂。我曾多次尝试想要掌握量子力学的数学运算，可最后都以失败告终，所以，面对围绕它展开的各种科学讨论，我只好充当无能为力的旁观者，尽管我也觉得有趣。但即使这样，对于博学多识的专家们所给出有关量子力学的诸多说法，我也仍持怀疑态度。

后来，我读了著名物理学家“夸克之父”默里·盖尔曼（Murray Gell-Man）的《夸克和美洲豹：在简单和复杂中冒险》（*The Quark and the Jaguar: Adventures in the Simple and the Complex*），这是一本专门为我们这些非专业人士撰写的科学读本。让我喜出望外的是，盖尔曼采用的是一种轻松的语调，字里行间透出启发性，他叩击混淆视听的歪门邪

说，透见了久违的光明，一举取得了大家的好评。请参照他“量子力学与胡说八道”（Quantum Mechanics and Flapdoodle）一章来理解我要表达的意思。我心里默念：“再来一击，盖尔曼，揍扁他们！”这位世界闻名的物理学家正在用我能明白的论述支持我对量子力学的种种偏见，这不正是我所理解的量子力学吗？但紧接着我又心头一颤，我真正理解了他的论证了吗？还是说，我只是近似理解了这些论证？我敢保证我没有受到他某种“花言巧语”的诱惑吗？此时，我不希望再有其他的物理学家出现，向我表明自己只是受了盖尔曼权威言论的蛊惑。我太喜欢他的结论了，不想再去考察他论证的细节。这跟你们的状况真是如出一辙。

但也有些不一样。我是在参照了很多专家的相关著作之后才对盖尔曼的作品给出了那样的评价，我的态度始终认真谨慎，至少到目前为止是这样的。而且，我也时刻准备接受，在之后的某一天，盖尔曼的那些“共识”也变成了想象力失败的一个案例而不再是之前深刻的观点，尽管它们是那么有魅力。我恳请你们也能用这种开明的态度来看待我在分解销毁塞尔中文屋的过程中所付出的那些努力，我也尽量把这件事情做得更易于让你们接受。

再回到20世纪50年代的图灵测试。在这个测试中，“法官”会试探性地跟两个实体展开交谈，一个是人，另一个是计算机。它们与“法官”对话时都要隐藏身份，只能通过“电传打字机”（可以想成屏幕配上键盘）才能完成交流。那个人要竭力向“法官”证明自己的真实身份，而计算机的任务则是要让“法官”相信自己才是人类，如果“法官”没能切实说出两者的不同，就证明计算机（程序）通过了图灵测试。这时我们说，计算机已经不只是具有智能，它“已经在最根本的意义上像人类那样具有了思想”（1988年塞尔在书中这样说道）。

对于这个领域的大多数人来说，为什么图灵测试会是检验强人工智能的试金石？因为和图灵一样，大家都认为，没有理解能力就完成不了那种对话，如果哪台计算机能成功完成这样的对话，就说明它已然具有了理解能力。[90]但事实上，图灵测试根本区分不出或者说不能区分哪个是僵尸哪个是“真正有意识的”人，因为一个有意识的人可以做的事情僵尸也照样能做。正如很多人所认为的，“通过图灵测试的实体，不仅是智能的，而且是有意识的”这个说法与僵尸直觉背道而驰。所以图灵测试不能算是个好的心智测试，但在仔细考察相关细节之前，我们先暂且保留这种判断。在我们形成一串意识流的地方，那个通过了图灵测试的程序就能形成一股殭尸式的非意识流，这明显对僵尸直觉构成了挑战，因为

僵尸直觉告诉我们这两者之间存在着实质性的差异。殭尸程序只拥有近似意识吗？让它功亏一篑的地方到底是哪儿？也许，塞尔的直觉泵就是想让所有贬低“强人工智能”的那些说法看起来都像是归谬论证。

塞尔请我们想象这样一幅场景：他被锁在一间屋子里，手动模拟一个巨大的人工智能程序，假设这个程序能理解中文。塞尔保证，这个程序能够从容应付图灵测试，能让自己经受住人类对话者企图将它从真正理解中文的人中区别出来的所有考验。而被锁在屋子里的那个人，塞尔，他不懂汉语，屋子里也没有其他理解中文的东西，塞尔能做的只是按照程序不停地操纵一些符号串，从中根本无法理解中文。屋里除了塞尔和“几张写着指令的纸条”，其他什么都没有。塞尔只靠精确地遵循这些指令手动模拟这个庞大的程序。如果塞尔不懂中文，那么当然（叮！），再加上几张纸他也照样不懂，所以，中文屋里没有可以对中文产生理解的东西，但它确实有强大的对话能力，足以在说中文的人中间以假乱真。塞尔和计算机一样，只会根据字符形状辨别出中文，对于他们来说，中文字符只是一些各不相同的无意义的“涂鸦”。所以，有塞尔在的中文屋只不过是计算机程序的一种实施方案而已，该程序所有的工作都无需对中文有任何理解，不管是在硅芯片上运行还是由塞尔操作。

真是个既简单又可靠的思想实验！还有哪里有错吗？哦，有。这个思想实验一在伯克利公开就引起了计算机专家们的反驳，塞尔将其称作“系统回应”理论，他们说道：

> 的确，锁在屋里的人类个体看不懂中文，但他只是整个系统的一部分，而整个系统是能形成理解能力的。个体面前摆着一本总账簿，上面写着各种规则，有大量的草稿纸和铅笔供他运算，他还有一整套中文字符的“数据库”。所以，形成理解绝不仅是一个个体的事情，它是隶属于系统的，个体只是其中的一部分。（Searle, 1980, p. 419）

塞尔随即形象有效地做出了下面的回应：

> 这让我觉得有些尴尬……，如此展开一个理论真的有些不可思议。你们想说的意思是，一个单独的人不能理解中文，但是以某种方式联合上几张纸条就能把中文弄明白？我真的不能想象，若没有某种意识形态的笼罩，一个人要怎样才能相信这

个想法。（Searle, 1980, p. 419）

让塞尔觉得不可想象的其实是图灵的那些基本洞见，利用它们图灵想出了有关存储程序计算机的理念！所有的能力都在软件那儿。回想在24章中我们介绍的寄存器机，单独的寄存器机根本不理解什么是算术，但是与软件联合起来，它们就能将算术完美实现。笔记本电脑上的国际象棋程序其实对国际象棋一无所知，但只要程序开始执行，它就肯定能赢过你，其他对峙也是如此，这就是笔记本电脑的强大威力。被塞尔称作意识形态的东西也是计算机科学研究的核心，它发展的每一步都走得坚定稳固。要完全复制出人类的能力并最终完成理解，需要在虚拟机上叠加虚拟机，在其上再叠加虚拟机，能量和能力完全由系统承载，它们从来不是隐藏在硬件里。图灵完美地重复了达尔文“奇怪的推理倒置”，创造出了自己的“奇怪的推理倒置”（Dennett，未发表）：在图灵之前，我们总以为人类所有的能力都来自于他们能理解，理解是所有智力的神秘源泉，现在我们才认识到，理解本身其实是一种效果，它从成堆的能力中冒出来，是各种能力层层叠加生成的。

细节是关键。塞尔从来没有告知他的读者，自己要在一个什么样的水平上手动模拟那个庞大的人工智能程序。他大概以为只要依照程序这样运行就可以：

> 现在，假设在接收到第一串中文字符（输入）之后我又接收到了第二串中文字符，与此同时还有关于怎样将两串文字联系起来的规则。规则是英文写成的，与英文母语的人们一样，我能够理解它们，能按照它的说明将一串规整的符号联结到另一串规整符号上。现在，还假设我又收到了第三串中文字符和一些提示，这些提示也是英文的，引导我将第三串字符中的某些因素与前两串字符联系起来，帮助我根据前面特定中文字符形状回应第三串中文字符中的某些相同形状的字符。（Searle, 1980, p. 418）

塞尔将如何联结这些“字符串”与“英文屋”做了对比：输入的不是中文而是英文句子或者故事，继而他直接用母语英文回应。

> 从屋外来看，即仅从某人阅读到我的“答案”这一层面来看，不管针对的那个问题是中文也好是英文也罢，我所给出的

> 中文答案和英文答案都不错。只是，对于中文案例来说，我是通过不经理解地操纵形式符号得到的，这与在英文案例中截然不同。（Searle, 1980, p.418）

看起来挺规整的一个类比！但我们要看看塞尔漏掉了什么。我们知道，塞尔收到了英文的“一套规则说明”，也就是所谓的提示，它们与“将寄存器39021192中的内容添加到寄存器215845085的内容中（机器码）”或者“定义一个常量的列队大小并将它设置在100（源代码）”这样的指令不同。塞尔是把自己闷在地下室，以每秒数万亿计次惊人的速度在操作计算吗？或者他是顺着程序的源代码在其上又叠加了多层操作？在遵循着源代码行进时他是否也读到了注释？最好不要，因为这些注释并非程序的一部分，它们只是为了给塞尔提示一些程序走向的线索，比如，“这是为了分析句子，生成名词、代词、动词、修饰语，并确定它是疑问句、陈述句、祈使句或者感叹句”，然后，在几十亿次操作之后，还会有“发现双关语，切换到应答模式……”，之后又是一套更加详尽的数十亿次的操作，在这个过程中，引用得到了精炼，备选答案就像对棋步不同走法的预测分析那样得到了评估，最后生成一个输出句子。

如果塞尔要手动模拟一个程序去进行一套感染人的中文对话，那么需要查阅数量庞大的数据库，可不仅仅只是“一系列中文字符”；正如他自己所说，至少要参考到以中文为母语的人能共享的所有日常知识。在塞尔手动操作的时候，他是根据提示进行各个层次的认知活动，还是完全仅依靠大量运算呢？

任何能够通过图灵测试的程序都必须经过一番“理智”操作，它与我们在展开对话时自己要经历的理智活动极为相似。假设，图灵测试中的提问者开始用苏格拉底式的问答方式裁选量子力学学科的候选人，给出一些简单的问题让这些测试者们来作答。要通过这样的测试，引擎室里的塞尔不得不通过精心设计的智力练习来阻止对话的结束，但当他结束这残酷的测试时，仍和刚进来时一样对量子力学一窍不通。相反，系统在测试结束时对量子力学的理解却深厚多了，因为它已经完成了相关练习。这个有关图灵测试的特殊案例会在程序中安装一个新的虚拟机，即简单的量子力学虚拟机。

我提到的这些事实，几乎全被塞尔那“几片纸张”和“规则”可以助他“关联”起中文字符的设想所掩埋。这并不是说，塞尔有意隐瞒了他所想象的手动模拟程序的复杂程度，我想说明的是，他忽视了这种复杂性

到底是一种怎样的复杂性。如果你所设想的程序只是相对简单的一系列规则，那么就像遗传学家贝特森思考“染色质颗粒”那样：“在任何已知的测试中都无法区分彼此，它们确实是同质的。”你当然会感叹，程序的理解力竟然那么“不可思议”，就像你同样也会感概DNA的力量不可思议一样。

现在看看，到目前为止我们都做了什么？我们已经转动了塞尔直觉泵上的旋钮，试图操控有关程序运行在不同层面上的描述。描述的层面当然有很多，在最高层次上，系统的理解力量没有超出我们的想象，我们甚至能洞察到系统如何理解了自己的所为。系统的回答不再令人感到尴尬，而令人感到很正常。但这并不表示，塞尔口中批判的人工智能已经达到能被称为有理解能力的那种水平，或者它们的运行方式在经由人工智能研究者想象延伸之后，就能够形成如此高超的能力。这仅表明，塞尔的思想实验并没有成功说明他想说明的那个问题：强人工智能不可能形成理解。

除此之外其实还有一些可以转动的旋钮，但我这个直觉泵已经在中文屋引出的庞大语境中得到了充分发挥。这里，我想集中讨论的并非中文屋所关注的理论和主张，而仅是“思考工具”本身，企图揭露出这件工具的缺陷，因为它在用晦暗不明的想象力而不是不断开拓出的正确想象力在说明问题。

61

从火星来的远程克隆人
他还是原来的那个“他”吗

你看到东方升起了月亮，西方也升起了月亮。你看到两个月亮在寒冷的黑色夜空上彼此靠近，紧接着擦肩而过，又继续在自己的轨道上行进。你站在火星上，距离自己的家园数百万公里，虽然有一层脆弱的薄膜（地球上的科技成果）裹着你，保护你在这颗红色星球的荒漠上免遭酷寒和死亡的威胁，但你已经坏了的太空船却怎么修都修不好。你被搁浅了，再也回不去了，你的星球、朋友、家人都永远留在了视线的尽头。

等等，也许还有希望。残破的太空船上有通信舱，你在里面找到了一个传送器“远距离克隆机马克四号”和它的使用说明书。你打开它，把光波调到地球上远距离克隆机可接收的波段，然后你进入发送舱，传送器会将你的身体快速分解，产生分子结构图，图纸通过光波发送到地球，随即，地球上接收舱的储备井里就填满了你身体所需的原子，它们会按照随光波传来的图纸迅速把你组装起来！你以光速返回了地球，你回到了朋友和家人的怀抱，开始跟他们细说你在火星上的奇遇历险。

最后一次检查太空船，你确定远距离克隆机是你最后的希望。走投无路的你打开了传送器，按下右面的按钮，进入发送舱。五、四、三、二、一，眼前一闪！你推开门，走出远程克隆机接收舱，来到阳光下，多么熟悉的气息，你回来了！从火星到地球那么远的距离，你被远程克隆了一下却依然毫发无伤。从那颗可怕的星球上死里逃生回来当然得开个庆祝会庆祝一下啦！朋友和家人都围绕在你身边，你突然发现，与上一次见到他们相比，每个人都有变化。毕竟都已经三年了，你也老了些。看看自己的女儿萨拉，她现在已经8岁了，你在想，“这还是当年那个坐在我膝上的小女孩吗？”“哦，这当然不假”，你马上反应过来，同时你也承认，要是单靠搜索记忆推断她的身份，自己可能真认不出她来。她长高了，俨然成了一个大姑娘，懂的东西也多了。事实上，她身上几

乎所有的细胞都已不是你当年看到的那些了。但尽管她长大了也变高了，尽管她身体的细胞早已更换过多次，但她仍旧还是你三年前吻别的那个小女孩。

马上你又心头一颤：“我果真还是三年前吻别过这个小女孩的那个人吗？我真的是这个8岁小孩子的妈妈吗？或者我已经是个全新的人了，只来到这世上几个小时，我只是保留了之前那些岁月所谓的记忆？”在远程克隆机发送舱里分解消失之后，这位年轻的妈妈是不是就死在火星上了？

我已经死在火星上了吗？哦，不，肯定没有，我这不是活生生地在地球上吗？有没有可能，有人，也就是萨拉的妈妈已经死在火星上了？那我就不是萨拉的妈妈了？但我肯定是呀！我走进那台远程克隆机就是为了能回到我的家乡、见到我的朋友和家人。但我现在有点儿记不起来了，是不是在火星上时我根本就没进过那台远程克隆机？如果真的有人进去过，是不是另有其人？

诡异的远程克隆机是运输工具，还是谋杀机器的一个代名词？萨拉的妈妈在这次与远程克隆机纠缠不清的经历中到底是生还是死？她想的是要活下去，她没有选择自我毁灭，她走进那个传送舱是为了有希望能生还。她冒险因为她无私，她必须采取措施，不能让自己的女儿萨拉失去那个深爱她想保护她的人；同时她也是自私的，为了脱离险境她只能殊死一搏。或者这些都只是看似如此。“我怎么知道这事情看起来会是怎样？因为我是当事人，是个心有所念的妈妈。我是萨拉的妈妈。事情就是这样。”

毫无疑问，我们可以远程传送一首歌、一首诗或是一部电影，这类东西“由信息构成”，那么，它们的本质在传送过程中会有缺损吗？要求可以在文档上留下电子扫描合法签名的呼声铺天盖地，想来，我们不情愿接受人的远程传送会不会像反对电子扫描签名一样有些不合时宜？哈佛大学学者协会是不会接受我推荐信上的扫描签名的，他们要求我用手的实际动作留下那串晾干的墨迹，所以我坐着出租车在贝鲁特转了大半天才拿到材料，在奶油色的高档书写纸上签了名，又连同相关表格给他们快递了过去。我了解到，现在学者协会改变了这项政策，但我仍然希望哈佛大学还是能在其颁发的学位证书上盖上封蜡章，尽管没什么利害关系，但这种传统承载着荣耀。

62

叙事重心
究竟什么是自我

什么是自我？哲学家们跟它纠缠了好多个世纪，这个关于“不朽灵魂”的概念，它非物质、不可言说，迷惑了众多思想家，让他们误入迷途，一下就是上千年。好在这个概念当年的势力如今已慢慢退去。“人死了，精神就会升入天堂”，人们越细想这个观念就越是觉得不合理，很明显，它的存在，连同着妖精和巫女的传说，只是为了不让我们抛弃这些念头。而我们之中的那些唯物主义者，他们认为精神其实就是我们的大脑，这种理解相对比较正确，但他们仍旧还是要面对这个问题：为什么人类会有一个看似是精神的东西？或许这样问更清楚：为什么每个人都有精神这么个东西进驻在身体里，更确切地说是大脑里呢？当我审视内心时，能找到那个自我吗？

著名哲学家大卫·休谟（David Hume）看不上这种想法，他在1739年写下了这段著名的言论：

> 就我而言，在我无比亲密地走近所谓的自我时，我总能无意中发现一些有关冷或热、亮或暗、爱或恨、痛苦或愉悦之类的独有知觉。只有动用知觉我才能捕捉到自我……如果哪个人在经过了严肃公正的反思后仍然认为，他心中的自我还有另外一副模样，那很抱歉，已经没有再辩论的必要了。我只接受，也许他与我一样有理，只不过在这方面我们是本质上不同的。也许他能清晰、连续地感觉到他所说的那个自我，但我在自己身上没找到这个东西。（1964, I, iv, sect. 6）

休谟半开玩笑似地承认，对于那些天真地从自身毫无创意的体验来推演别人经历的人来说，如果他们总是用这样的方法来思考“我是不是个僵尸（当然也包括殭尸）”，那么从别人那里得到的可能总是不同的回声。我这么猜的，可别当真。

什么不是自我倒是清晰明了。自我不像杏仁体或者海马体，它不是大脑的一部分。大脑额叶在我们考量自己的境况、目的、知觉等时总是起着至关重要的作用，即使这样，也没有人会错把自我定位在大脑额叶那里。前额叶切除术是一种很可怕的外科手术，它“只为患者的自我留下了一个影子”，但它并不因此就变成了一场自我切除手术。就像一个很老的笑话讲得那样，我宁愿在那儿放只空瓶子也不愿做那个前额叶切除术，不过话说回来，不管哪种情况，“我”都实实在在经历着。

那自我会是什么呢？我觉得它应该是一种类似于重心的东西。重心是一个抽象概念，尽管它很抽象，但与物质世界紧密结合在一起。就像其他所有的物质事物一样，你也有一个重心，准确说是质量中心，不过在此我们先忽略这些细节问题。如果你的重心比同体重的人高的话，那么保持直立对你来说就比别人要难一些。有很多方法可以测出重心，它的位置取决于很多因素，比如穿的鞋子或者最近一次吃饭的时间，随着这些因素的变化，重心在身体中间的一小块区域游走。它是一个理论上的点，不是哪个原子或者分子。一根钢管的重心并不是由钢构成的，也不由任何什么东西构成，它是空间中的一点，在贯穿钢管中心的中间线上，到钢管两端的距离相等。

仅就重心这个概念来说，它就是一件不错的思考工具。重心实际是把事物的物质颗粒联合地球上所有物质颗粒所受的引力平均化后得出的，它最终归结到了地球的中心（地球的重心）和事物的重心这两个点上，能帮助我们计算出事物在不同条件下的行为。举个例子，一个物体，只要它的重心超出了支撑架的托举面，就肯定会掉下来。当然，在牛顿发现地心引力这个概念前，我们对重心就早有一些直观的理解，比如“快坐下，船摇得厉害”。

现在我们来说说这个概念在细节上是怎么操作的，以及它为什么这样操作。举个例子，如果我们打算设计一辆车或者一台落地灯，要想降低它们的重心，或者使重心处于更合适的位置，在好多设计因素上，我们就必须得考虑“重心”这个概念了。也许它只是“理论家的一个虚构”，但那也是一个非常有价值的虚构，从它上面能生出很多真实的解释。如此这么一个抽象的、没有物质存在的东西真能引出些什么吗？并不那么直接，但是将使用重心的解释与单纯使用因果关系的解释比较一下，重心的作用就体现出来了。为什么即使帆船倾斜得厉害，咖啡杯也不会翻倒？请比较一下“因为咖啡杯的重心很低”和“因为咖啡杯紧贴甲板”这两种解释。

我们会把重心看成是理论家的虚构，原因可能是它其中掺杂了一些虚构色彩，即有关不确定性的一些奇怪设定。阿瑟·柯南·道尔在自己写的推理小说中也赋予了福尔摩斯很多特点，他没有写到的就说明福尔摩斯不具有。作者从没有提到过福尔摩斯有三个鼻孔，所以我们就有理由认为福尔摩斯真的没有（Lewis, 1978），我们也知道他从没有重婚。对于某些问题，我们大可以如此推测，然而，还有这样一些问题，它们是无解的。在他的左肩上是不是有颗痣？他是不是英国作家王尔德的堂弟？他在苏格兰有自己的别墅吗？对于一个真实的人来说，类似这样的成千上万个问题，即使我们不去询问，答案也会安静地待在那里。但于福尔摩斯而言就不一样了，他只有柯南·道尔为他虚构出的那些特点，以及在此基础上推论出的一些东西。一个单纯的读者可能会相信福尔摩斯确有其人，他也许很想知道，在开往奥尔德肖特的火车上，那个售票员的个头比福尔摩斯高还是矮，但如果明白什么是虚构，他就不会再问出这样的问题。这种情况也适用于重心这个概念：如果你已经把“重心”看成是理论家的虚构，还卯足了劲儿一定要弄清楚“它到底是不是中微子”，那么我只能说你还未得其要领。

叙事重心又是什么呢？它同样也是理论家的虚构。设置这样一个概念是为了能把构成一个人的那些变幻莫测又纷繁复杂的动作、表达、烦闷、抱怨和诺言等统一起来，让它们变得可以理解。它负责组织在人类层面上给出的各种说明。你的手签订不了一个合同，而你能；你的嘴巴说不出谎言，而你可以；你的脑子记不住巴黎，而你行。你是你鲜活身体“所有履历的所有者”，根据它我们才知道你是你。就像我们所说，你的身体投你所好。我们可以把世界所有其他的地方和树立于地面的方尖塔之间的万有引力简化成两个点间的作用力，也就是地球的中心点和方尖塔重心间的作用力；同样，我们也可以把两个自我之间的互动，例如握手、对话、墨水涂鸦等简化成是买方与卖方，两者只是在完成一笔交易。每个自我都是一个人，有自己的履历、“背景”，还有许多正在做的事。与重心不同的是，自我没有贯穿着时间和空间的轨迹，他一边行进一边收集自己，随时积攒记忆，一路设想着自己的预期和计划。

每个人都想否认掉自己身上的一些不光彩的地方，但做了就是做了，已经展开的叙事无法再修改，除非借助传记的力量在之后重新诠释。“做那件事的时候，我已经完全不是我自己了”，我们常说这句话，而能容忍如此矛盾的自我辩解也算是我们的明智之举。要是一个人在做过一件事之后马上对你说“这绝对不是我想做的”，那他想表达的可能是，恳求你不要对他当时的行为妄下定论，不要从这些事情推断他未来

的作为。你可能会相信，但也可能不信。类似这样的说法还有：“我不去做，谁做？”“我鬼使神差地做了那件事。”重申一下，对于这些话，你通常接受的不是表面意思，这些话真正否定的是指导说话人行动的那些性格和动机，这会牵扯到我们下一章中要阐述的两个概念“责任感”和“自由意志”。但是，现在请注意，我们需要借用自我这个概念在你做了什么和什么事情发生在你身上之间划出一条界线，先不论这条界线划得有多武断。

每个物体都有一个重心，每个活生生的人也都有一个自我，更确切地说，每个现存人类的身体都是某个自我的身体，自我住在身体里，是身体的管理者。而这样的管理者只有一个吗？一副身体就不能由两个或者多个自我共同拥有吗？像多重人格疾患“表面上”就是多个自我进驻一个身体的情况，其中有一个主导自我（那个“主人”），后面跟着一群“替代者”。我之所以会加上“表面上”三个字，是因为业界其实对这种病情的诊断一直存在争议，大家从那种明目张胆的欺诈行为一直争论到二联性精神病（在这种病例中，医生如果经验不足，稍不谨慎就会让病人的精神变得更加混乱），再到一些很罕见又真实存在的病例，这些病例往往有很多人添油加醋地模仿。多年来，我和心理学家尼古拉斯·汉弗莱也一直在探讨这些研究和处理方法，后来我们得出结论——所有人都是对的！欺骗和夸大其词随处可见，也总能碰上热切的病人对峙好骗的医生，如此种种。哦，是的，尽管这些走火入魔的谈话者都有他们详细的说辞，但其中的很多状况还是能在一些最基本的案例中找到。我们毫不惊讶地认识到，多重人格只不过是正常生活状态的加剧版本而已，我们多多少少都体验过。我们大部分人都在工作上、家庭中过着截然不同的生活，其中每一种情境都会让我们养成不同的习惯和记忆，但这些习惯和记忆并不能很好地适应其他情境。

就像社会学家欧文·戈夫曼（Erving Goffman，1959）在他有关这个话题的一本经典著作《日常生活中的自我呈现》（*The Presentation of Self in Everyday Life*）中所叙述的那样，我们总在把自己展现成真人戏剧中的某个角色（丹尼特教授，从路人到爸爸再到爷爷等等），也能轻而易举地得到配角们的配合，因为他们也在以同样的方式展现着自己。为了完美地实现自我展示，我们互相串通，随时准备为彼此的生活轨迹添砖加瓦。如果有谁因为怠于扮演本来的角色而破坏了剧情的顺利延伸，那结果就会变得尴尬、好笑，或者更糟。忠于职守需要钢铁般的意志。想象一下，你刚在宴会上碰见一个人，他要怎样展现才能表明自己的身份？是一本驾照，一本护照，还是一头向你冲过来，准备给你一个

大大的拥抱？人在极端的困境中，往往会采取比较极端的应对措施，而假装已深陷绝望几乎成了他们的第二天性。当事情变得越来越棘手，你径直走开，只留下另一个不同的叙事重心、一个不同的角色，它要全副武装，随时准备应对麻烦。

我和汉弗莱认为，在采访一个假定患有分离性身份识别障碍的患者时，你总觉着他会是一位在绕过强行询问技术上炉火纯青的大师。为了看出性格的一面是否真的会对性格的另一面正在做什么或者说什么没有任何记忆，你要提出一些问题要求对方立即作答[91]，或最好设置一些小陷阱，为此你必须彻底变得粗鲁，甚至严重冒犯对方；你还会发现，对方总能礼貌地戏弄过去，连眼都不眨。行骗高手们能凭借高超技术有意为之，但人格障碍的无辜受害者在转换性格时却并不对此有所意识。在某种程度上，性格转变在每个人身上都会发生。但这只是虚构的人物，不是吗？真正的人格才是主人，对吧？然而，事情并不那么简单。

事实是，那个你扮演的人，不管是多重角色还是单一角色，他才是你的叙事重心。你的朋友通过你的扮演角色认识你，“你今天不像平时的你”，在多数情况下，你自己也得通过这种方式看待自己，当然这稍显夸张了，“哦，我的天呐！是我做的？我不会这么做！”专业小说家和大骗子手一样，总能巧妙地运用叙述手法把一些细节扩大化。而我们这些其余的人，虽然不专攻写作但也是才华横溢，基本上，我们不知不觉就会围绕那些细节巧妙地缠绕上我们的故事，像蜘蛛织网一样，完全是自然为之而绝非艺术行为。与其说我们在用大脑编织我们的故事，还不如说是大脑在用故事编织我们自己。当然，每个人都有一个实实在在不可否认的核心经历，然而，随着岁月流逝，其中的大部分会渐渐变得无效，形同虚设，与你现在是谁再无多大关系。在你的生命历程中，为了自我维护和自我改善，你可能还会主动放弃、否定、甚至“忘掉”自己的某些经历。[92]

有些关于你过去的事情很容易回答。你是否跟一位电影明星跳过舞？你去过巴黎吗？你骑过骆驼吗？你曾经徒手把人掐死过？答案或许不一定，但回答起来还算简单。试想一下去问某个人最后一个问题，他突然停了下来，陷入沉思，然后开始抓耳挠腮。如果出现这种情况，你要注意跟他保持安全距离！其中的原因是，我们对自己足够了解，知道自己是否曾跟一个电影明星跳过舞，是否去过巴黎、骑过骆驼，或者掐死过某人。

现在开始回忆往事，如果问题中的事情“没有出现在脑海”，我们就

把这种空白理解成对提出的问题给出否定回答。为什么我们如此肯定这就是正确回答的方法？你会为所有从未做过的事情列个清单吗？你会为所有从没去过的地方列个清单吗？再来看看下面这些看来差不多的问题：你跟一个叫史密斯的人跳过舞吗？你去过一个卖地板蜡的杂货店吗？你乘坐过一辆蓝色的雪佛兰吗？你是不是打碎过一只白色的咖啡杯？这些问题，有的很好回答，有的你无法回答。有的问题太琐碎了，你甚至回答错了都不自知。为什么你能记得自己是否做过这些事情呢？很多我们经历的事情其实并不怎么让人难忘，有些我们做过的事情，无论好与坏早已从我们的叙事重心上脱落掉，也有很多没有发生的事却傻傻地被我们的叙事重心吸收了进来，原因是它们可能有助于我们成为人物t。你所拥有的是经验和才华、庄严的意图和白日梦的幻想，把它们的总和装进一个大脑和身体里，赋予一个特定的名字，那就是你。认为你、自我精神中除此之外还有一种特殊的不可分割的东西，这只是一种吸引人的幻想。要想去了解一个人，问问他的理想、希望、英雄主义或者罪恶就足够了。

叙事重心虽说不是一块有关神秘精神的稀世珍宝，但毕竟也是一种抽象的概念，如此，我们是否能够系统地研究它？是的，可以。

63

他者现象学
人类语言的妙用

他者现象学（heterophenomenology）并不是直觉泵，它是一个脚手架，我们有必要在解决一些棘手的问题之前将它搭建起来。意识研究涉及现象，乍一看，这些现象似乎发生在另一个维度：即私有的、主观的“第一人称”维度，每个人都有自己的意识，而其他人无法直接访问。那么，在研究流星、磁铁或人体代谢、骨密度这些标准的“第三人称”客观方法与研究人类意识的方法之间有什么关系呢？我们是不是需要开创一些彻底的、革命性的替代科学？或者，让我们对这种标准方法加以扩展，使之能够公允地对待意识现象？我捍卫以下主张：存在一个对客观科学简单而保守的扩展，它能宽宏地涵盖人类意识的所有基础，公平地对待所有数据，而且，并不需要放弃已经在科学的其他领域中运行良好的实验方法的约束条件与规则。这种第三人称方法便是他者现象学，一种关于他者而不是自我的现象学，这种方法可以尽可能严肃合理地采用第一人称视角。

为什么会有这么一个诘屈聱牙的名字？“现象学”（phenomenology）原指一系列林林总总的现象，它存在于一个完善的科学理论出现之前。16世纪，英国物理学家威廉·吉尔伯特（William　Gilbert）编撰了一部出色的现象学书籍，收录了形形色色的磁现象，而在几个世纪之后，他精心描述的那些磁现象才全部都得到了解释。20世纪早期，埃德蒙德·胡塞尔（Edmund　Husserl）以及一批深受其影响的心理学家和哲学家，采用了“现象学”（Phenomenology，首字母“P”大写）这一术语，用来称呼一种关于主体经验现象的科学研究，它用“第一人称”内省的方式进行观察，试图保持理论中立从而避免任何理论上的预设。这一思想一直延续到了今天，但其大部分内容早已四面楚歌、被人们忽视，原因有好有坏。尽管它带来了一些非常诱人、值得进一步探索的成果，但这种第一人称的方式已经被客观的经验主义

的科学所遗弃，后者着意强调数据，因为所有研究者都可自由获取它。但是，我们可以客观地研究意识，这种方法实际上是对现象学的扭曲，因此，我把它称作他者现象学，以区别胡塞尔式的自我现象学。他者现象学是一种基于客观科学的第三人称视角来研究第一人称现象的现象学。

显然，在对岩石、玫瑰、老鼠所做的实验和对意识清醒、能够合作的人类主体所做的实验之间有着重要的差别，后者能用语言交流，并因此可以同实验人员沟通，他们通过言语互动提出建议并报告他们在各种可控实验条件下是什么样子。这是他者现象学的核心：它利用我们执行和解释言语行为的能力，产生了一系列主体关于其意识体验的信念。这一系列信念充实了主体有关他者现象的世界：一个取决于主体S的世界，一个主体的主观性世界。他者现象学的全部细节，外加我们所能收集的同时发生于主体大脑之中以及周遭环境里的所有事件，构成了人类意识理论的全部数据，我们必须对此予以解释。这里面没有遗漏任何有关意识的客观现象和主观现象。

该解释需要采用意向立场将声音、按键行为等实验中的原始数据转变为对信念的表达和报告：它需要一个工作假设，即在考虑主体的历史感知和需求的前提下，假设主体是一个行为人，其行为受到信念和欲望的理性指引，而这些东西本身是理性的。例如，在标准防范措施中，意向立场的各种约束可以清楚地在这样的实验中识别出来，这便可以防止主体已有的经验使其在报告他们的信念或欲望时出现偏差，从而扭曲了我们对其行为的解释。我们并不希望他们知道我们期望他们报告什么，因此，与此同时，我们需要采取措施使他们理解我们布置给他们的任务。采用意向立场并不是完全主观和相对的。解释的规则能够得到清晰地说明；不同主体对于解释达成一致的标准能够制定出来；偏差可以得到确认；对理性所做的不可避免的假设也能得到谨慎地对待，从而可以让我们将其视为一个可调节、可辩护并在进化上可解释的假设。（Dennett, 1991a）

这里并不是要提出一个研究意识的新方法。我只是发现认知心理学、心理物理学（研究物理刺激和主观反应之间的关系）和神经科学领域的研究者已经采纳了这些标准方法，并自觉地解释、捍卫了这些方法。只要我们正确地理解并遵循这些方法，就能避免任何激进的“第一人称”意识科学革命，并且不会为可控的科学研究留下任何无法研究的意识现象。

这一方法论主张要求我们怎么做呢？所有的科学都致力于做毫无疑问的事情，例如神经元和电子、时钟和显微镜，而与此相比，该主张要求更高，它要求我们只是全力关注信念和欲望：这些信念由主体所表达，它们被视为主体性的基本要素；欲望指的是同实验人员合作的欲望，并尽可能坦率地向他们报告真相。这种方法的一个重要部分包括对这些信念和欲望的控制，如果有实验结果表明无法做出这种控制时，我们就必须放弃这一实验。那么，信念和欲望是什么呢？对此，我们尽量不予表态，等待理论上的确认，但我们可以将信念连同它们的内容、对象视作理论工作者的虚构和抽象，就像质心、赤道或力的平行四边形法则这类东西。

美人鱼目击报告是真实的事件，但该报告内容总是不实的，因为美人鱼并不存在。同样的道理，一组关于体验的信念并不等同于体验本身。哲学家约瑟夫·莱文（Joseph Levine, 1994, p. 117）就此表示反对，“意识体验本身并不仅仅是我们的言语判断，它是理论必须回答的原始数据。”这不可能是正确的。我们要怎么做才能先于理论就将体验本身分门别类？设想我们将主体置于实验环境中并对其进行询问，同时要求他们做出我们所要求的行为，借此我们可以取得证据。这些证据资源被我们所进行的级联式解释自然地套嵌在了一起，在这里，我将它们按其粗糙程度排列如下：

“意识体验本身”

关于这些意识体验的信念

莱文提到的“言语判断”

表达言语判断的某种表述

在“原始”意义上，表述是原始数据，它记录了声音与动作。如果条件允许的话，脑电图、功能性磁共振成像以及此类装置的读数也可被算做原始数据。基于可靠的解释方法我们可以接受（3）和（4），据此我们就有了一系列的主体信念，在这些条件下，这些信念是怎么样的？但是我们是否应该事先推进（1）？这不是一个好想法，有两个原因：

首先，如果（1）成立而（2）不成立，即，如果你有意识体验，而你自己并不相信，那么这些额外的意识体验对你来说是无法理解的，你只是一个外在的观察者。

因此，相比于他者现象学，莱文提出的这些个替代物并没有提供更多的可用数据。

> 其次，如果（2）成立而（1）不成立，即，如果你相信你拥有意识体验，而事实上你并没有这种意识体验，那么这种信念就需要解释。

如果我们坚守他者现象学的标准，将（2）视为原始数据的最大集合，并且避免对任何可疑的数据做出承诺，同时确保任何人都可以对所有现象做出理解。

在言语判断中，如果有一些信念是难以用言语表达的，那它们会是什么？他者现象学家和主体可以毫无限制地进行合作，通过设计模拟或其他非语言的方式表达此类信念。例如，

> 在下面的这条线段上画一条垂直线，以表明在某个维度上有多强烈的体验：
>
> 几乎无法察觉————非常显著

或者，主体可以以不同力道按压一个按钮，以此表示疼痛、焦虑、无聊，甚至对试验的疑惑程度。然后，我们就可以用大量依赖于生理的变量，即从皮电反应、心率，到面部表情以及体姿的变化来衡量主体的意识体验。作为一个主体，如果你在用尽了这些方法之后，依然相信残留着无法表达的东西，那么，你可以将它报告给他者现象学家，他们可以帮你将这种信念添加到你的原始数据的信念列表之中：

> S宣称他对X有着无法言喻的信念。

如果这种信念是真实的，那么科学有责任解释此类信念是什么，以及它们为什么会无法言喻。但若这种信念是虚假的，科学依然得解释S为什么会错误地相信存在这种特殊的、无法言喻的信念。[93]

64

色彩科学家玛丽：揭露吊杆托架

错误假设导出错误结论

澳大利亚哲学家弗兰克·杰克逊（Frank Jackson）关于色彩科学家玛丽的思想实验通常被看作是“知识论证”，自1982年发表以来，就一直以其非凡的活力不断引发出哲学家的直觉泵。单就其直觉泵的数量和可靠性而言，它必须被算作有史以来一个分析哲学家所发明的最成功的直觉泵之一。它简直就是一个经典，被永久地列入心灵哲学本科课程的必读书目，几本很有分量的选集也专门讨论过它的含义。有意思的是，它的作者随后宣布放弃思想实验的结论，但这并没有削弱它的影响力。

下面便是这一思想实验的全部内容，也许略有变动，不过依然非常清晰：

> 玛丽是一位天才科学家，不知出于什么原因，她被迫待在一间只有黑白色的屋子里，并通过一台黑白电视研究世界。她专攻视觉神经生理学，让我们假设，当我们看到一个熟透了的西红柿或天空，并使用诸如“红色”“蓝色”等词汇时，她熟知此时出现的所有物理信息。比如说她发现，所有的过程不过是来自天空不同波长的光波组合刺激到了视网膜，随后，该刺激通过中枢神经系统引起声带收缩，并将空气从肺部排出，最终，我们说出了这句话：“天空是蓝色的。”不可否认，原则上，我们可以从黑白电视机上获得所有这些物理信息，否则，函授大学必然会用到彩色电视机。当玛丽走出黑白屋子，或是给她配上一台彩色电视时，会发什么呢？她能学到什么？结论似乎相当明显，她会学到有关这个世界的知识以及我们对这个世界的视觉经验。不可避免的是，她以前的知识是不完整的。但她不是早已熟知了所有的物理信息吗？因此，存在着比这一切更多的东西，所以，物理主义（即唯物主义，该理论否认二元论）是错的。（Jackson, 1982, p. 130）

这是一个优秀的直觉泵吗？让我们来转动旋钮，看看它是如何工作的。实际上，这可能会花费我们非常长的时间去熟知相关文献，不过，这项工作我已经做了。因此，在这里，我将简单地阐明一些有待检视的旋钮，把结论留给你们当作练习。如果乐意的话，你可以得出自己的结论，你能想出一个新的转动旋钮的方式吗？ 20多年前，我曾对这些旋钮做了初步的分析，并发表了一个令人扫兴的结论，结果，它被人们严重忽视了："对于一个优秀的直觉泵，缺乏经验的人很难一下子就看出其要点。然而，事实上，它是一个低劣的思想实验，它其实是在蛊惑我们误解它的前提！"（Dennett, 1991a, p. 398）

让我们看看是否如此。我认为，想象这样一个场景要比人们通常假定的那样来得更为困难，他们想象的东西太过简单，总是从一个错误的假设出发得到了他们的结论。

> 第一个旋钮："待在一间黑白屋里通过黑白电视研究世界。"

假设她戴着黑色或白色的手套，不许她在洗澡时看自己的身体，但是，切断色彩的所有"外部来源"这种想法在任何情况下都是无法实现的。难道我们要给她装上某种装置以防她的眼睛看到她自己或者试着给她造成"光幻视"（phosphenes）？难道在她看到实际色彩之前，她就不能在梦中看到？如果不可以的话，为什么会这样？难道在她将色彩"存储"于大脑之前，必须"通过眼睛获取"？在这个简单的假设背后，有关色彩的民间理论乱作一团。

> 第二个旋钮："当我们看到一个熟透了的西红柿或天空，并使用诸如'红色''蓝色'等词汇时，她熟知此时出现的所有物理信息。"

出现的所有物理信息？有多少呢？是不是就像拥有世界上所有的金钱那样？那会是什么呢？要想象这一点并不容易，其困难程度不亚于想象所有那些使该思想实验可以得到它想要得到的结论的东西。它必须包括发生在所有大脑中的有关反应的全部信息，包括玛丽的大脑，尤其是，它必须包含任何条件下对所有色彩的全部感情以及情绪反应。因此，玛丽得非常详细地分辨，哪些色彩会让她感到平静，哪些会使她心烦意乱，接触哪些色彩会加深她的印象，哪些色彩会让她心不在焉，而

哪些又会使她厌恶等等。她被禁止用自己做实验了么？这可既没有作弊，也没有将任何彩色的东西偷偷地放进她的小屋子。如果你无法想象所有的这一切或更多，那么你就无法跟进这一思想实验。这就像是被要求将一个圆设想成一个1 000边形。从一个心理表征（mental representation）的练习能推导出很多含意，而这些含意并不能由另一个心理表征推导而出。比方说，在这个案例中，我们是不是应该忽略这样一个事实？即，如果玛丽熟知所有信息，那么，毫无疑问，由于担负着成千上万的百科全书词条和图表，她整个人将处在一个心力衰竭的状态。

如果杰克逊曾假定玛丽拥有“物理全知”的能力，不只是关于色彩，还关于每个层面上所有的物理事实，小到夸克，大至星系。所有读者，就算不是全部，也都会反对，认为这样的壮举太不切实际了，很难当真。但假定玛丽仅仅知道关于色觉的所有物理事实基本上不是幻想。

> “让我们想象玛丽有10亿个脑袋……”
>
> “别犯傻啦！”
>
> “好吧。她有100万个脑袋……”
>
> “就这么定了！当真？”

现在我相对简单地将故事变得富有戏剧化，我鼓励人们思考一个完全不同的结局：

> 有一天，把玛丽关起来的人觉得是时候让她看到色彩了。他们使了一个小花招：准备了一个鲜艳的蓝色香蕉作为玛丽的第一次色彩体验。玛丽看了看说：“嘿！还想骗我！香蕉是黄色的，但这东西是蓝色的！”关她的人大吃一惊。她是如何做到的？“很简单，”她清楚，“要知道，我可是知道一切的一切——迄今为止所有已知的关于色觉的物理因果。所以，毫无疑问，在你把香蕉拿给我之前，我早就记下了一个黄色或蓝色物体，甚至一个绿色的对象等，会带给我的所有物理感受，这些感受会对我的神经系统产生影响。所以，我已经准确地知道了我会产生什么样的感受，毕竟，做出这种思考的‘纯粹性’不就是你那赫赫有名的‘感受质’吗？我一点儿也不对我的蓝色感受感到惊讶，使我感到惊讶的是，你居然会使这种下三滥的把戏。我

> 知道，你很难想象我会对自己的知道得如此之多，以至于我对蓝色的感受丝毫不感到惊讶。当然，你很难想象这一切。任何人都很难想象有人知道有关物理一切的一切！”（Dennett, 1991a, pp. 399-400）

人们大都认为事情不会这样发展下去。杰克逊漫不经心地说道，“实在太明显了，她会学习到有关这个世界的知识，以及我们对这个世界的视觉体验。”乔治·格雷厄姆和特里·霍根（Graham & Horgam 2000, p. 72）也会说道，“当然（叮！），我们觉得，她应该既感到惊讶又觉得高兴。”但这是一个错误，这正是玛丽这个思想实验的错误所在。当玛丽第一次看到色彩时会像得到了某种启示，这个结论因为看上去实在不错，所以之后便没人再愿意费心去思考这个故事该如何发展了。其实，它根本就不该如此编排。

杰克逊的直觉泵极好地暴露了大量关于色彩体验本质的天真思考，毫无疑问，大脑在大部分时间都为我所有，所以，我们承认，他很好地从民间理论那儿引出了一些有内涵的东西。但他的目的是要反对这种假设，即物理科学有能力解释所有的色彩现象。当然，在真实世界中，处在玛丽位置上的人会学到一些新东西，因为，关于色彩，无论她知道多少，总是会存在大量关于色彩的物理效应，对此她并无知晓。该思想实验只是一个被设想得很极端的案例，它使杰克逊的“实在太明显了”以及格雷厄姆和霍根的“当然”不得其所。如果你依然倾向于认为我对故事结局的改动根本是不可能的，那么，且看看你能为你的信念提供什么样的辩护。如果你的一番考量能避开数以百计的哲学家这么多年来艰苦卓绝的思索，那倒是件非常有意思的事情。当然，它毕竟是一个无与伦比的直觉泵，30年来，它让哲学家们“有口饭吃”。

第六部分小结

大家都觉得自己是行家！这种现象严重困扰着有关意识的系统研究。当然，情况也没有那么糟糕，我只是觉得，有些人只对某个问题思量过几分钟就认为自己想出的观点比所有根据高科技实验结果以及大量数据得出的结论还要有道理，这实在是有点儿说不过去。在很多科学研讨会的问答环节，每每听到他们又引用一些近期自认为有力的体验去更正自己在演讲者的工作中发现的错误时，我就觉得有点儿可笑。如果如这些人所想，只有人类自己才能对自身经历的所有特性给出最可靠的评判，那他们倒真是没错！

但是，受一些不可靠却极有说服力的意识形态的控制，我们也会记错，或是错误地描述，甚至误解我们最亲近的那些体验。有这么一个例子，你在家里就可以完成，它很简单，却可能让你很震惊。为了按照下列指示观察自己，请先在镜子前面坐下。现在，开始专心凝视镜中自己的眼睛，不要受外界干扰，就像面对着靶子一样注视它们。视线保持不动，一只手摸索着从一打洗好的纸牌里抽出一张，把它举在离你一臂左右但脱离你视线范围的地方，牌面朝你，晃动纸牌。显然，你肯定知道你在晃动它，但是你却看不到它。现在慢慢将纸牌移入你的视线，同时保持它在摇动，这时，你会看到它在摇动，但你却看不清上面的图案！你说不出纸牌上是红色、是黑色、还是一张人面牌，更不用说上面的数字了。纸牌渐渐向你的注视点靠拢，此时，你更要集中精力，不能作弊，不能偷偷看纸牌。直到你真正可以看清楚纸牌的时候，纸牌几乎就在你的正前方了。

惊讶吧？我想，在第一次体验到这一点的时候，肯定每个人都很惊讶。我们一直认为，视线“从中心集中点到外周”，对细节和色彩的关注度应该都是一样的，现在我们明白了，尽管这种想法看起来“合乎情理”，也能经得住一般推敲，但它终究是错误的。“（视觉）意识呈现给我们的是一个丰富、连续、细腻的世界”，这只是我们的幻觉。所以，在科学界还没有给出最新发现之前，先不要急于对所见意识现象妄下结论。对此深表不屑的一些哲学家在摇椅上想出的那些理论才根本微不足道，那些理论更多带来的是困扰和迷惑。我们“通过反省”而“了解”的意

识只是意识中次要的一些性质，但它们却在我们利用他者现象学框架思考意识并对意识展开系统性研究时，极严重地误导了其中的很多思考。

仍然会有很多难题有待分晓，但此难题非彼困难问题。如果在解决完所有简单问题之后还是剩下一些不可知的神秘残余，那我们就该重新考虑下自己的出发点了，有时候，得用一些激烈的方式才能让我们从当代生物、物理，甚至逻辑学的假设中脱离开来。借此机会还能看看，利用常规科学我们到底能走多远，尽管人类目前对所有事物的现行理解都是常规科学带来的，不管是行星、板块构造论，还是生物的繁殖、生长、修复和代谢。

第七部分

关于自由意志的思考工具

INTUITION PUMPS
AND OTHER
TOOLS FOR THINKING

在思考这个主题时，我们很容易被常识映像与科学映像之间的差异所欺骗。就像面对“颜色是什么？颜色实际上是什么？美元是什么？美元实际上是什么？”等问题一样，在我们开始思索“自由意志到底是一场幻觉还是我们真正拥有的东西？”时，我们实际上是希望从科学映像入手来探索这个问题，但这个问题本身却是从描述常识映像的传统表达中提出的。近年来，用科学映像探索这一问题的热情越来

越高涨。有一批相当杰出的科学家直截了当地说过，自由意志就是一种幻觉，比如：神经科学家沃尔夫·辛格（Wolf Singer）、克里斯·弗里斯（Chris Frith）和帕特里克·格格德（Patrick Haggard），心理学家保罗·布卢姆（Paul Bloom）[94]和丹尼尔·韦格纳（Daniel Wegner），以及一些大名鼎鼎的物理学家，比如斯蒂芬·霍金和阿尔伯特·爱因斯坦。难道这么多伟大的科学家都错了？许多哲学家，虽然不是所有或许也不是大多数哲学家会说："是的。这是哲学的工作！"他们的回答正确吗？我想是的。

科学家通常会犯一种低级错误：把常识映像与人们对常识映像的流俗看法混为一谈。坦白地说，人们对颜色的流俗看法荒谬之极；颜色可不是大多数人所想的那样，但这并不意味着常识世界中真的没有颜色，我们只能说真正的颜色与人们所设想的颜色有相当大的差别。人们对意识的流俗看法也是荒谬的，完全是一种神秘主义二元论。如果意识果真如他们所想，那么赖特就是对的（见第59章）："我们不得不承认，意识并不存在。"可我们其实不必把意识当作由神奇组织所构成的"真正的魔法"，这种东西并不存在。我们需要首先承认人们流俗看法中的意识是不健全的，然后才能认识到真正的意识现象。同样，自由意志也不是人们流俗看法所宣称的那样，是一种跳出因果性的神奇现象，我把它比拟为一种悬浮。还有一位为这种奇谈怪论做辩护的哲学家大言不惭地宣称，每一次自由选择都是一个"小奇迹"。如果科学家们异口同声地将这种自由意志称作幻觉，我百分之百赞同。但这并不意味着在与道德有关的任何一个意义上，自由意志都是幻象。它像颜色一样真实，像美元一样真实。

然而，有些科学家在宣称科学已经证明自由意志是一种幻觉之后，又在道德意义上继续强调这个"发现"。他们认为这一发现将对道德与法律产生重要影响，因为再也没人能够真正对自己的行为负责任了，所以我们再也不应该惩罚或者赞扬谁了。就像有人会说：没有固体，至少没有真正的固体。这些科学家们也犯了人们常犯的错误。他们所用的是过时的、流俗的概念。就像我们在处理颜色和意识，还有空间、时间、固体性等其他很多这些常识映像时，人们的流俗看法是错的，这是首先应该改变的。

这一部分介绍的直觉泵是为了让你从对自由意志的流俗看法中解脱出来而设计的，它们会让你看到一种更好、更真实、更实用的概念，看到这一常识映像的要旨所在。几千年来，人们围绕自由意志这

一主题争论不休，相关内容已经过于庞杂，以至于任何一章甚至一本书的体量都无法将它说清楚。但是我们不得不从某个地方开始。我们的起点就是思考工具，它们的作用就像铁锹，把陷入老车辙中的你解救出来，助你驶向新的天地，从一个更好的角度看待这一问题。为了说明这项工作有多重要，我设计了第一个直觉泵。

65

一位恶毒的神经外科医生
非手术的手术

现在我们正处在神经外科诊疗领域的黎明，对心理状况的考虑在神经疾病的治疗中已经被大大弱化了。比如，神经精神病学家达米安·德尼斯（Damiaan Denys）和他的同事们所做的开创性研究（Denys, 2010）显示，在强迫障碍的治疗中，通过埋入电极进行深层脑部刺激可以产生惊人的疗效。这是事实。但接下来就是我的虚构了：有一天，一位出色的神经外科医生，在闪闪发亮的高科技手术室里给她的患者做完植入手术后，就对患者说：

> 我给您植入的装置不仅能够治疗您的强迫障碍，还能控制您的每一个决定。这要感谢我们的主控系统，它可以与您脑中植入的微型芯片随时联络。换句话说，我已经解除了您的自由意志，以后您对自由意志的感觉都只是幻觉了。

事实上，她没有这么干。这只是一句谎言，她想欺骗她的患者，看看会发生什么。谎言奏效了，这个可怜的家伙出院以后，果真相信自己不再是一个能对自己行为负责的行动者，而只是一个傀儡，他的行为也体现出了这一点：他开始变得不负责任、好斗、随随便便，纵容自己最坏的念头肆意妄为，直到被送上了法庭。在为自己辩护时，他激动地宣称自己不应该为那些行为负责，因为他的大脑被人植入了芯片。神经外科医生来到法庭作证，她承认自己说过那些话，不过她补充道：“但我只不过是在跟患者开玩笑，这只是一个恶作剧而已。谁能想到他真会相信我的话呀！”

无论法庭最终采纳了谁的证词，把谁送进监狱，都不重要。神经外科医生那段欠考虑的玩笑话已经毁了这个患者的生活，夺去了他的正直，严重伤害了他的决断力。事实上，她对患者做的假“通报”，反而以一种非手术的方式完成了她声称自己要通过手术完成的事：她让患者变

成了“残废”。如果她要对这个可怕的后果负责的话，那么某些神经科学家也危险了，现如今，媒体上充斥着他们的言论，说科学已经告诉我们，自由意志只是一种幻觉。这些人不也正以同样的方式在伤害那些对他们的话信以为真的人吗？[95]神经科学家、心理学家和哲学家务必严肃对待自己的道德义务，在对这些问题发表公开声明之前，请仔细想想这些话的前提和含义，就像我们要求人们必须先想想清楚再对全球变暖或者小行星撞击地球的问题大加评论一样。让我们看看犀利的社会批评家和观察家汤姆·沃尔夫在科学家的声明中找到的这些话吧：

> 实验室之外的人所描画的结论是这样：我们的大脑是电路固定好的！无法改变！别怪我了！是我的电路接错了！（Wolfe, 2002, p. 100）

电路接错了？可是接对了又怎样呢？如果科学家们“发现”，我们谁也没接对，或者根本不可能接对，那么道德义务又该如何呢？

66

生命游戏
我们的世界是被决定的吗

每当物理学家理查德·费曼听到别人谈论他所不熟悉的科学领域时，他总爱问这么一个问题："关于你所讲的这些，能给我举一个简单的例子吗？"如果说话人没能满足这个要求，费曼就开始怀疑了，这理当如此。这个人当真有什么要说的吗？还是在空谈些花哨的术语以炫耀他在科学上的才智？如果你无法相对简化一个困难的问题，就可能意味着你没有找到解决这个问题的正确方法。简化并不只是新手的事。

生物学中的"模式生物"是科学家们精心选择的一些物种，可以让他们的实验相对简化。这些物种能够在实验室中迅速繁殖，相对安全且容易处理，只要对它们进行分组研究就能很容易地做出比对、得出结论。这类生物有：果蝇、小白鼠、斑马鱼、鱿鱼（因为它有着巨大的神经轴突）、秀丽隐杆线虫、拟南芥（一种抗寒的、生长迅速的芥菜类植物，它是第一个完成完整的基因组测序的植物）。人工智能研究中也有一些简化的案例，比较有名的是"玩具问题"，顾名思义，它是对真实世界中那些"严肃"问题的简化。人工智能专家们设计的很多极有意思的程序都是在解决"玩具问题"，比如，有一个程序是在积木世界搭建出简单的结构，整个虚拟世界是由一个平面和一堆积木块构成的。下国际象棋也是一种"玩具问题"。确实，这比开车从缅因州到加利福尼亚州、解决巴以冲突、甚至在厨房里做三明治都更省事。伦理学家则有电车命题，说个最简单的版本：一辆失控的电车行驶在铁轨上，若不变道它将撞死铁轨上的五个人；你手里拿着一个开关，它可以让电车变到另一条轨道上，而如果它冲上另一条轨道，将会撞死一个人。那么你会按下开关吗？

下面我要介绍一个可以帮助人们思考决定论的"玩具世界"，决定论的意思是：某一时刻的事实，包括每个粒子的位置、质量、方向、速度等会决定下一时刻发生什么，以此类推，接连不断。物理学家和哲学家花了上千年的时间来争论我们的世界是不是决定性的，是否有一些真正

的未确定的事件，比如没有任何原因所导致的完全不可预测的“随机”事件的发生。即使是有经验的思想家，也会在生活的游戏中找到新的见解。1970年，数学家约翰·何顿·康威（John Horton Conway）和他的研究生们创造了一个让人叹为观止的决定论世界的简化模型。

生命游戏在一个二维网格中发生，就像一个棋盘，可以用鹅卵石或者硬币充当棋子；当然也可以高端一点，用电脑屏幕来显示这个二维网格。这不是一个有胜负之分的游戏，如果硬说是一种游戏的话，它更像是单人跳棋。网格把平面分成了一个个小方格，每个方格的状态要么是关、要么是开。如果是开，在格子中放一枚硬币；如果是关，则让格子空着。我们可以看到，每个方格周围都有8个方格：东、西、南、北4个相邻的方格与东北、西北、东南、西南4个对角的方格（见图1）。

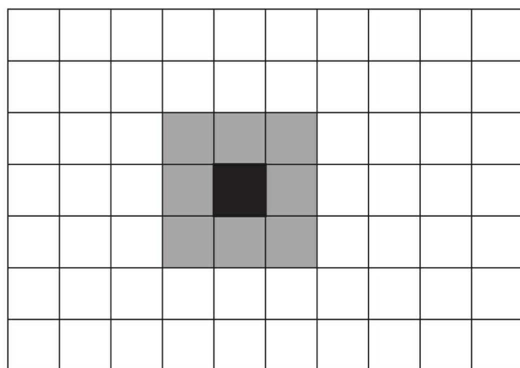


图1

在生命游戏的世界中时间是离散的，而不是连续的。时间滴答滴答地前进，游戏世界依照如下规律在滴答之间发生变化。

> 生命物理：对于每个方格，我们都要计算在当前时刻与它相邻的8个格子中有几个状态是开，几个状态是关。如果一个方格周围有两个格子的状态是开，那么在下一时刻，该方格将保持现有的状态（开或关）。如果一个方格周围有3个格子的状态是开，那么在下一时刻该方格的状态为开，不论它现有的状态是什么。如果是其他的情况，那么该方格在下一时刻的状态是关。

这是生命游戏的唯一规则。现在你已经知道关于生命游戏的全部规则了，生命游戏世界中的物理学就体现在这个简单且普通的规则中。既

然它们是这个生命游戏世界中基本的物理定律，那么我们首先就可以借助定律从生物学角度来理解这个神奇的物理现象：把小方格的开当作生，关当作死，时刻的变动当作世代的交替。不论是过度与世隔绝（即周围生存的方格少于2），还是过度拥挤（周围生存的方格大于3）都会导致死亡。下面让我们思考一种简单的情况。

如图2所示，现在只有格子D与格子F拥有3个处在“开”状态的相邻格，所以只有它们会诞生新的一代。而对于格子B与格子H而言，它们都只拥有一个处在“开”状态的相邻格，所以它们的下一代将会死亡。而格子E有两个处在开状态的相邻格，所以它会保持不变。那么在下一时刻，情况将如图3所示。

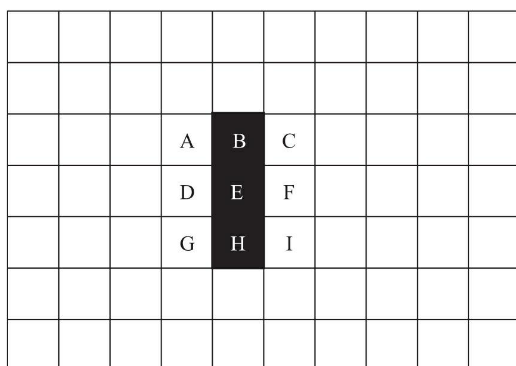


图2

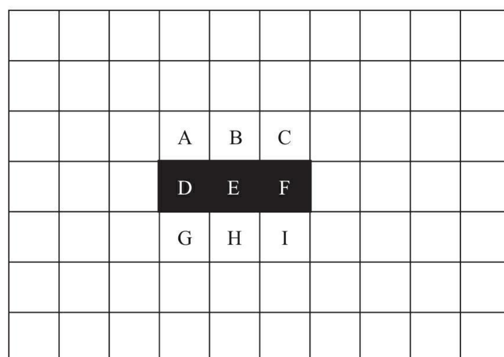


图3

很显然，该图形会在下一时刻还原到图2的模式。除非场面上突然多了一个状态为“开”的方格，否则这两个图案的交替模式就会啪嗒、啪嗒一直反复不停地进行下去。我们把它称作闪光灯或者信号灯。下面请看图4，下面的图形将会如何变化呢？

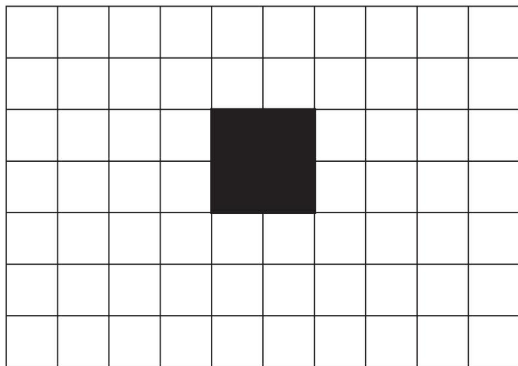


图4

完全不变！因为每个“开”的方格周围都有3个“开”的格子，所以它们会继续保持原有的样子。而每个“关”的方格周围都不存在3个“开”的格子，所以也不会有新生。我们管这个图案叫作静止的生命。

只要谨慎地遵循这个唯一的定律，面对任何图形，人们都可以准确地预测出下个时刻、下下个时刻、下下下个时刻……哪些格子开着、哪些格子关着。可以说，19世纪初期著名法国科学家皮埃尔·拉普拉斯（Pierre Laplace）所设想的决定论在生命游戏这个“玩具世界”中得到了完美的展示：给出世界在某一时刻的状态，我们这些观察者就可以利用物理定律完美地预测出世界在未来任何时刻的状态。或者我们也可以这样说：只要我们站在物理的角度，就可以完美地预测生命世界，没有误差，没有不确定性，只有唯一一种可能。并且，这个二维的生命世界还告诉我们，没有观察不到的东西。没有后台，没有隐藏变量，事物的物理现象在生命世界中直接展露，完全可见。

如果你觉得用这个定律一遍遍地演算非常枯燥，也可以使用电脑来模拟。你可以在电脑程序里输入初始图形，让它来帮你执行运算，图形就会根据那个唯一的规律一次又一次地改变形状。你可以利用更好的模拟程序随意调整时间和空间的跨度，时而聚焦局部时而俯视全局。

人们很快就能发现，有一些简单图形会比其他图形更有意思。让我们设想一条斜线，如图5所示。

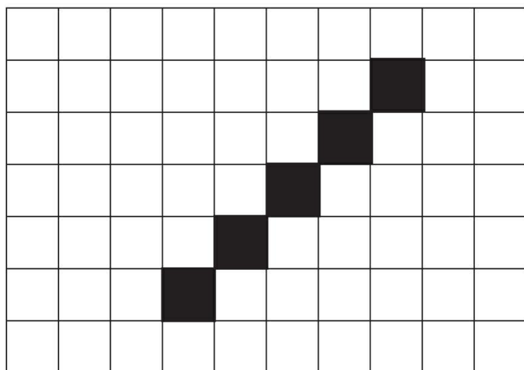


图5

这个图案不是闪光标。每一次更新，该斜线两端的两个“开”着的方格都会因孤立无援而死亡，而且也不会有新的诞生。过一会儿，整条斜线就消失了。除了永远不变的图形（如“静止的生命”）和逐渐消失的图形（如这条斜线）之外，还有很多可以进行周期性变化的图形。比如我们已经见过的闪光标，它形成了一个两代的周期，除非有其他的图形侵入，否则这种周期性变化会永远持续下去。侵入现象是为生命游戏增添趣味的一个元素。在进行周期性变化的图形中，有些可以像变形虫一样游动。其中最简单的是滑翔机，这个由5个像素组成的简单图形看起来就像是在往东南方飞去，如图6所示。

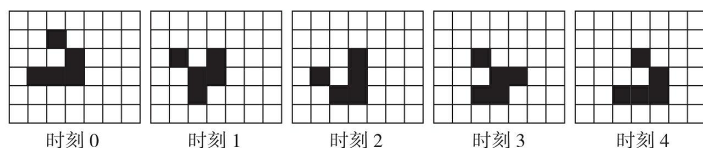


图6

除此之外，生命游戏世界中的“居民”还有很多，比如吞噬者、河豚列车、空间之耙以及其他一些名字起得很妙的家伙们，以这样的方式命名意味着我们是在一个新的、类似于设计的层面上认识它们。与在物理层面上对它们作出枯燥的描述相比，使用这个层面的语言显然为我们省去了不少麻烦。比如，美国作家威廉·庞德斯通（William Poundstone）曾这样来描述图7：

> 吞噬者将在四个时刻之后吃掉滑翔机。不论吞噬者捕杀什么，基本过程都是一样的：首先，吞噬者和它的猎物形成接触。在下一刻，双方接触的部分都会因过度拥挤而死亡。随

后，吞噬者进行自我修复，而它的猎物不会。只要猎物残存的部分继续向前滑翔并且自我消亡，那么就算吞噬成功。（Poundstone, 1985, p. 38）

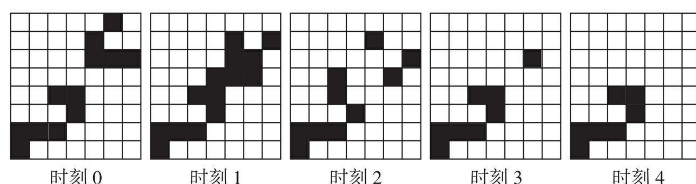


图7

请注意，当我们转换到一个新层面上来谈论这些图形的时候，我们在“本体论”上，即我们有关存在的总目录上也会发生些微妙的变化。在物理层面上，这个世界没有移动，只有开和关，唯一存在的个体是那些有固定位置的小方格。而在设计的层面上，我们一下就“看到了”移动的持续存在对象。在图5中，向东南方移动并不断变形的是同一架滑翔机，尽管在每一时刻它都是由不同的小方格构成的，而在图7中，吞噬者“吃掉”了一架滑翔机，也就是说生命世界里少了一架滑翔机。

我们还应该注意，因为在物理层面上只有普遍的定律而没有例外，所以我们在设计层面上的概括必须滴水不漏：我们概括时必须加上“通常”或者“假如没有别的东西侵入”等等限定语。上一刻留下的一小块杂物，也许在下一刻就能将这个本体论层面上的“东西”“摧毁”或者“杀死”。对这个层面来说，可以把小方格当作具体物来看待的这一特性相当重要，但却不那么保险。这一特性之所以重要，是因为人们上升到这一层面，采用这种本体论视角，才能够预测出更大图形构造或者结构系统的行为动向。这样做会冒一点风险，但却省去了在物理层面的繁琐计算。比如说，我们一个人自己就能设计出一些有趣的超级系统，它们是脱离“部分”而言的，我们可以试着利用一些在设计的层面上可用的“零件”来设计一个超大的图案系统。

康威和他的学生们就给自己设定了这样一个目标，并且完成得极为出色。他们设计并证明了：在生命游戏中创造出一个能够自我复制的物体是可能的。他们设计的东西可以完美地复制出一个新的自己，而复制品会继续复制，直至无穷，无可阻挡地扫过这无尽的平面。另外，它还是一架通用图灵机。原则上说，它可以运算一切能在计算机上运算的函数！

康威和他的学生们到底受何启发才创造出了这样一个世界，并在其中创造出了如此神奇的东西呢？他们试图在一个非常抽象的层面上回答生物学的这个核心问题：一个可以自我复制的东西所需的最小复杂性是多少？他们遵循的是约翰·冯·诺依曼在很早以前想到的一条巧妙思路，冯·诺依曼直到1957年去世前一直都在钻研这个问题。诺贝尔生理学或医学奖获得者弗朗西斯·克里克和詹姆斯·沃森在1953年发现了DNA的结构，但DNA如何工作却一直是一个谜。冯·诺依曼非常详细地设想过一种漂浮机器人，它可以利用一些零散的碎片来制造自己的复制体，并不断重复这一过程。他还描述了一个自动机是如何读取它自身的设计图、并将其复制到它创造的那个新个体中去的。他很具体地预见到了后来所发现的DNA表达和复制机制。为了让自动机在数学上既严格又易于处理，冯·诺依曼把问题转换到了简单而抽象的二维平面当中，这就是著名的细胞自动机。生命游戏中的那些小方格就是细胞自动机的一个恰当实例。

康威和他的学生们想要具体地证明冯·诺依曼的理论，他们打算创造出一个具有简单物理定律的二维世界，并在这个世界中创造出一个可以稳定地进行自我复制的装置。像冯·诺依曼一样，他们希望自己的答案尽可能地一般化，这样才能最大限度地独立于现行的（地球的？本地的？）物理学与化学。他们需要的是极为简单的东西，既容易想象又容易计算，所以他们不仅把问题从三维世界转换到二维世界，并且还把时间与空间数字化了。就我们看到的，时间和空间完全由“时刻”与“方格”的数量来代替。冯·诺依曼曾在图灵机设想的基础上，详细设计出了通用存储程序串行处理计算机，现在我们把它称作冯·诺依曼机，还精彩探讨了这样一个计算机所需要的空间和结构特性。冯·诺依曼已经意识到，在原则上我们可以在二维世界中制造一台图灵机，并对此给出了证明。[96]而康威和他的学生们也打算在他们设计的二维世界中证实这一点。[97]

这可不是一件容易的事，但他们还是向我们展示出了如何利用一些简单的生命形式“构造”出一台可工作的计算机。比如，一串运动着的滑翔机可以变成输入或输出的“纸带”，而吞噬者、滑翔机和其他零零碎碎的小东西可以组合成一个读写器。那么，这台机器看起来是怎样的？庞德斯通经计算得出，整台机器将占用大约10^{13}个小方格或像素：

> 要想显示一个10^{13}像素的图形，需要一个至少有300万像素宽的显示屏……想象一个高分辨率的屏幕，类似你的笔记本电脑或iPad，不过它的宽度要达到800米。如果从全景上看，

这个自我复制图案中的像素会小得几乎不可见。当你距离屏幕足够远，让整个图案完全进入视野中，其中的一个个像素会变得极为微小，即使它们组合成的滑翔机、吞噬者或一把手枪也无法辨识。整个自我复制的图案构成一条朦胧的光带，看起来像是银河。（Poundstone, 1985, pp. 227-288）

换句话说，当你在二维世界中用足够的零件去组合一个能够进行自我复制的东西时，这些零件就像是原子，而那个东西则像由原子组成的微生物，两者之间的大小实在相差太多。而且你或许也无法让那个东西的复杂度降得更低了，虽然这一点没有经过严格证明。

生命的游戏可以阐释很多重要的原则，我们可以用它来构想许多不同的论证或思想实验，但在这里我只想指出其中三点，其他的留给你自己去发掘。第一，在生命游戏中，物理层面和设计层面之间的差异是怎样变得模糊起来的。比如说，滑翔机到底算是一个被设计出来的东西还是一个像原子和分子一样的自然物？如果真要说的话，康威和他的学生们用滑翔机、吞噬者之类的组件构造出来的纸带阅读机才肯定算得上是被设计出来的东西，而那些小零件只是原材料，是生命游戏世界中的简单“物”。滑翔机无需设计发明出来，它只是我们在生命游戏的物理学预设中发掘出来的。当然，那个世界中的一切东西也都是如此。在生命游戏的世界中，所有东西都是严格地蕴含在它的物理学和那无穷无尽小方格的构造中，通过那套简单的定律我们可以直接从逻辑上推出这一结论。只不过，在那个世界中有某些东西看上去比别的更加奇妙，更加令理解力低下的我们无法预料而已。有些人认为，康威的那台可以进行自我复制的计算机，也就是这条像素银河也“仅仅”是一个生物大分子，只不过其运动周期极长、极复杂。这倒是很好地解释了一些人在生物学和生命的起源问题上的一个平行看法。有人会说，氨基酸就是氨基酸，它无需被设计，可是仅凭氨基酸就能合成蛋白质，这简直太神奇了，至少得经过点儿什么近似设计吧。看来，还需要我再讲一遍达尔文的进化论啊。

第二，生命游戏的世界是决定论的，我们可以预测所有可能图案的未来，但奇妙的是，它们的过去往往完全是个谜！想想由四个开着的方格组成的静止生命吧。不管是观察它还是它的相邻方格，你都无法从中获悉它过去的状态。说到这一点，请注意，只要那四个方格中有三个开着，那么在下一刻它们就都将是开着的，“静止的生命”就会出现。而过去到底哪一个方格是关着的，这是一个无效的历史事实。

第三，回想一下“噪声”和碰撞对于突变的发生有多么重要，与其他创造性的过程一样，进化也源于突变。可是康威构建的巨大的自我复制体却无法发生突变。它总是完美地复制自身，若要把突变机制引入进来，整个构造必须增大数倍。为什么呢？因为生命游戏的世界是决定论的，唯一可能的“随机”突变只能由一些游荡的小物体带来，它们貌似随机地在屏幕上游动，破坏别的物体。但是，在生命游戏中能动的最小物体就是“滑翔机”了，你可以把它想成是单光子或者宇宙射线，以生命物理学中的光速在运动。其实单单一架“滑翔机”就能造成很大的破坏，如果非要让它只是对一个自我复制体的基因组织作出“调整”而不是摧毁这个基因组，那么这个基因组必须比“滑翔机”巨大得多，也要足够强壮。如果事实证明，那些银河般巨大的集合体实在脆弱，无法经受住一阵偶然降下的“滑翔机雨”的考验，或许就可以很好地说明，无论我们在生命游戏的世界中设计出的物体有多大，它都不会发生进化。

67

石头、剪子、布
最好的套路就是没有套路

大概你们每个人都知道石头、剪子、布的游戏。两个人面对面，说着“石头、剪子、布”，然后同时伸手。石头崩剪子、剪子剪布、布包石头。除非双方手势一样，否则总能分出胜负。这是一个吸引人的游戏，因为如果你能看透你的对手、猜透她的心思、伸出对的手型，你就可以一直赢下去。但若真的发生了这种事，也是怪吓人的。难道真的有比别人更会猜拳的人吗？显然，我们早已举行过奖金丰厚的猜拳锦标赛，也产生出了国内、国际的猜拳优胜者。而且，参加锦标赛就一定得有人赢，即便其中根本没有技巧可言。更好的选手就是一直赢下来的人。

他们是怎么办到的？或许他们可以从对手的表情和姿势中读出微妙的变化。扑克玩家称这为阅读对手的“tell”，分辨他们是不是在虚张声势，同时自己还要绷着一张“扑克脸”。也许我们在玩石头、剪子、布时也会流露出“tell”，一般人很难控制，所以最好的猜拳者可以在出拳前最后一刻捕捉到它。那么，为了不让对手从你的外在行为中找到规律，最好的策略是什么呢？最好的策略就是绝对随机地玩猜拳，因为只要你一直随机地出，你的对手也就无迹可寻了。在你随机出拳甚至破坏性出拳时，你还可以顺便观察对手的猜拳模式，利用这一模式现编一套战术，然后不再随机猜拳，给对手一个突然袭击。

众所周知，实际上人类并不善于创造随机的序列，他们倾向于变换动作，避免把同样的动作重复做两三次。而在一个完全随机的序列中，这种情况本来是相当常见的。因为知道自己随意编排的无规律序列是很容易失败的，所以你打算找个更好的办法，比如从图书馆或者网上找一张随机数字表。“随机”地选择表上的某个数字，把它之后的上百个数字抄下来。首先要把所有的0删去，然后把所有的1、2、3用代表石头（rock）的R替代，所有4、5、6用代表布（paper）的P替代，所有7、8、9用代表剪子（scissors）的S替代。这样，你就能够得到一个包含大

约90次猜拳的随机序列，因为你删去了大约10个0，这对于比赛来说应该够用了。

你已经准备妥当，可以比赛了，现在最需要恪守的原则是不要把这张表告诉任何人。一旦让你的对手看到了它，你就等着任人宰割吧，就像俗话说的，她会把你当成摇钱树的。而另一方面，只要对手得不到你的“出招表”，她就不得不费劲地猜，绞尽脑汁地预测你的一系列动作。简单地说，她将不得不站在意向立场上理性地分析你，不得不推理你为什么那么出；而一旦让她看了你的出招表，她只要依照这张表就能预测你的出招了。

“不要让对手看到你偷偷选择的是什么”，这条简单的原则实际上触及了人们争论不休的自由意志问题的核心。事实上，冯·诺依曼和经济学家奥斯卡·摩根斯坦（Oskar Morgenstern）能发明博弈论，就是因为他们开始意识到：虽然我们可以通过收集信息，然后用概率论算出效用值的方式预测出单一行动者或意向系统未来的行为，但如果面对两个行动者或者两个意向系统，情况就会产生根本性的变化。现在，两个行动者都必须考虑到对方也会预测自己的行为，他们必须把对方的观察、预测、算计也纳入到对自己行为的考量中，这会造成一种异常复杂而不确定的反馈循环。(98)

从根本上来说，在每个行动者看来，对方都笼罩在迷雾般的不可理解和不可预测性之下，这一信念为博弈论的兴起创造了条件。人们在对生物进化的研究中也发现了这种博弈现象，许多物种间的交互影响看上去都像是使用了博弈论的原则（无需理解的能力！）。瞪羚的弹跃只是一个简单的例子。还有蝴蝶难以捉摸的飞行，这样它就可以让食虫鸟很难预测它的飞行轨迹，即便鸟类通过进化获得了比人类更快的“闪光融合率”（flicker fusion rate）。鸟类能看到比我们更多的“帧频”，对它们来说，看电影就像看幻灯片。

让自己变得不可预测，同时小心别人也这样做！我们发现，动物们普遍都在本能上“领会”了这一原则，当它们面对任何一种活动复杂的存在时，都会将其看作一个行动者。为了安全起见，它们会“问”“来者何人，你想要什么？”而不只是问“那是什么？”，因为那个东西可能会把它吃掉、和它交配或者与它争夺猎物。这种本能反应也是人们发明的那些看不见摸不着的事物的进化源泉，它们从哥布林、矮妖、精灵、食人魔和众神，最终进化成神，进化成一个终极的看不见摸不着的意向系统（Dennett, 2006a）。

像动物们无需知道自己为什么能采取这种策略一样，我们人类也可以领会这一技巧，保持自己的不可预测性而无需理解这一策略的好处在哪儿。而在很多时候，它的好处的确非常明显。当你在古玩市场看上一件古董时，你明白最好不要先对它夸赞一番再去跟老板谈价，那样你会被“宰”得很惨。而当你要卖一个东西时，最好开价合理，开一个令你满意的价格，因为你不能指望有买家会出高价来买你的东西，你应该避免的是让买家了解你能接受的最低价格，当然你也可以拒绝讨价还价。在拍卖中也是，如果你把报价提前告诉了拍卖师，那么你只能指望他足够正直，不把你的最高价透露给其他的竞标者。

同样的，如果你一见钟情，死心塌地爱上了某个人，也最好不要一下子意乱神迷、脸红心跳，你要尽可能地保持克制、平常对待。否则你会把你的梦中情人吓跑的，即便她没跑，这也会让她在感情中占尽上风。“扑克脸”不只是为扑克玩家准备的。一般而言，如果你的竞争者或你周围的其他行动者对你毫无戒备，你会更容易获得自己想要的东西。他们预料不到你想干什么，也就无法对你做出防备。时刻提防他人是很费神的，所以除非你的竞争者掌握了确切的证据，否则他们根本不会去试图预测你要干什么。

魔术师们知道如何利用“心理力量”使你“出于自己的自由意志”选择那张他们希望你选中的牌。这类招数很多，全都非常微妙、令人难以察觉，而一个真正优秀的魔术师总是能够做到这一点。实际上，这是对你的主动性的剥夺，魔术师是在把你变成他本人意志的延伸，变成一个工具、一个棋子，使你不再是一个自由的行动者。

与古代的自由观不同，我们不再向往完全独立于因果性之外的那种自由选择。我们所向往的，或者说我们应该向往的自由观是：当我们行动时，可以充分了解到我们力所能及的最好选择是什么。如果我们关于“那里究竟有什么？”的真信念都是由周围环境导致的，那么周围环境也一定导致了我们根据那些信念做出最明智的行动！作为一个行动者，关于周围环境的信息几乎就是我们想要的一切，但是，我们一定不希望周围环境中冒出一个反操纵你的行动者。所以，我们不希望自己的意图表现得太明显，不希望自己的所思所想让别人知道太多，因为他们可能会利用这一点。所以，我们的愿望清单中还应该加上保持自己思想和决定的神秘性的能力，即便这意味着，为了防止他人对我们做出防备，我们有时只能选择次好方案。国际关系学者利亚姆·克莱格（Liam Clegg）在2012年对此做出了一项具有开创性的形式分析。

有些人糊里糊涂地执着于不可预测性，他们认为“为了安全起见”，我们应该坚持一种绝对的不可预测性。要想达到这种不可预测性，要求我们的大脑运行的物质基础在物理层面上是非决定论的。哲学家杰里·福多尔曾生动活泼地阐述过这一思想：

> 人希望自己成为传说中那偷吃禁果的夏娃，可以完全自由地做任何事。实际上，她是如此自由，就连上帝都不知道她会去偷吃禁果。

但是，为什么“人们希望”这样的自由？绝对的不可预测性真的比实践的不可预测性还要好？几千年来，无数哲学家坚持认为绝对的不可预测性是自由意志的先决条件。难道他们了解一些我们所不知道的事情吗？如果真是这样，他们一定保密得很好。对于他们中的大多数人而言，“自由意志与决定论不相容”这一命题俨然无需论证。但是要我说，他们有论证的责任。这些哲学家们应该告诉我们，为什么没了绝对的不可预测性我们就应当感到绝望。而我已经证明过了，为什么尽量保持一定的实践的不可预测性对我们而言是明智的，就像进化本身。所以，请告诉我们吧，为什么这还不够？

我倒是替他们找了一个好理由：如果你打算和上帝玩石头、剪子、布而且赌注颇高的话，比如赌你会不会得到拯救，那你肯定会像福多尔所说，希望拥有“完美的”自由。而我个人，并不期待这样一场赌局，所以我很满意我所拥有的实践的自由。如果我要赌个大的，只需要做好保密工作并且离那些神偷和操盘手们远一点就好了。

68

两种彩票
从新视角看“决定论”

公平地比较一下两种彩票。代表着“之后”（after）的彩票A：彩票销售结束后，所有的号码被放到摇奖器里摇奖（要多随机就有多随机），最后抽出中奖号码。我们玩的很多彩票都是这样的。另一种是代表着“之前”（before）的彩票B：在彩票销售之前，中奖号码就已经通过摇奖的方式确定好并被保存到一个保险箱里，而在其他任何方面，它都与彩票A相同。有人可能会觉得彩票B是不公平的，因为在人们购买彩票之前中奖号码就已经决定了。在这些彩票当中已经有了一个胜者，虽然没人知道是哪一个，所以其他彩票都成了不值钱的废纸，把它们卖给不知情的民众是一种欺诈行为。可实际上两种彩票是同样公平的。每个买彩票的人都有着相同的中奖概率，这与中奖号码在何时选出没有任何本质上的联系。

之所以很多彩票都把摇奖环节放在了最后，是为了让公众作为直接见证人，证明彩票摇奖没有作弊。这样做可以避免掌握内部消息的人偷偷将中奖结果泄密，因为在彩票售完之前，根本不存在也不可能存在知道中奖号码的内部人士。有趣的是，并非所有的彩票都是这样运作的。“出版商结算所”（Publishers' Clearing House）每年都会寄出上百万封信，每封信上都用粗体字写着“您可能已经中奖了”，奖品是一百万美元或者其他什么奖品，现在已大多改作网上发放。他们之所以打出这么昂贵的广告，就是因为有市场调查显示，人们普遍认为只要程序上是公平的，事先选好中奖号码的彩票也可以是公平的。但是也有可能，人们之所以毫无怨言地接受了这种彩票是因为他们不用花钱。

如果我们从一开始就知道中奖号码已被密封到专门的信封里、存进了银行的保险柜，还会有很多人来买这种彩票吗？有的，有成百上千万人在买“刮刮乐”，在人们购买时，无论中没中奖都已经是事先决定好的。显然人们认为他们真的有机会中奖。我想他们是对的，不管他们有

没有道理，是否对这种中奖过程的公正度有信心。这应该会扰乱一些哲学家的心神，这些哲学家莫名其妙地让自己相信，除非在最后一刻才决定结果，否则这些都不算真的中奖机会。他们坚持认为，如果不能持续不断地提供真正的随机性，没有随时出现不确定的分支点去打破因果关系的大网，那么自由选择将是不可能的，我们将没有真正的机会去做正确的事。

这两种彩票可以让我们从一个新的角度看待决定论问题。如果世界是被决定的，那么我们拥有的只是一种虚假的随机数字发生器，而不是真正的（量子力学的）随机性发生器。如果世界是被决定的，那么我们所有的彩票实际上在大约140亿年以前就已经开奖了，随着宇宙大爆炸的一声巨响它们就被装进了信封，并在我们需要它们的时刻交到了我们手上。每当你要“抛硬币”或者不怎么夸张地做一个随机的决定时，你的大脑就会把那个信封打开，从中取出一个“随机”数字，然后你会让它来决定你该干什么，就像我们在“石头、剪子、布”那一章中讲的那张“出招表”一样。

“但这不公平啊，”有人会说，“有些人分到的胜利比其他人多呀。”确实如此，在一切具体的分配活动中总有人摸到的牌比其他人的好，尽管人们应该谨记，从长远的角度上看，运气是平均分配的。但是如果在我们出生之前所有的签就已经被抽好了，那么有些人就会注定比别人幸运！说的没错，但是，即便在我们出生前签还没有抽好，即便它是随着我们的生活进程遇事现抽的，结果会有什么不一样吗？即便抽签机制完全随机、公正，是纯粹非决定的，也仍然有些人注定比其他人赢得多。因为即使在一个完全公平、随机的比赛中，比如掷硬币，也注定有某个人会赢，而其他参赛者会输。获胜者可能不会说这是“命中注定”，可不论我们把他的胜利归于什么，是命运、是他自己或是别的什么，难道会有人说这不公平吗？公平并不在于所有人都赢。

人们期望非决定论的最常见的理由是：如果世界是决定论的，那么只要我们选择了某个行为，就意味着“我们不再可能有别的做法”，但是，“能有别的做法”对我们来说当然（吖！）应该是一件很重要的事。同样，这一点也没有它看起来那样理所当然。为了让大家看到这种想法会对我们产生的误导，下面请思考一类有趣的无效历史事实。

69

无效历史事实
用强行分类抹平微小差异

无效历史事实是一些非常平常的事实，它发生在过去，只不过现在人们已经无法对它做出判断了，这是一种在当今世界根本没有留下任何痕迹的事实。我最爱举的一个无效历史事实的例子是：

> A．我牙齿里面有些金子曾经属于凯撒。
>
> B．我牙齿里面的金子曾经并不属于凯撒。

根据逻辑，以上两句话中必定有一个是事实。等等，根据“近似算子”那章的精神，我们不应该怀疑这些“显而易见”的选言式命题吗？让我们想想，怎么可能A和B都没对事实有明确的表达呢？如果凯撒时代的所有权是不明确的，那么凯撒只会近似拥有一部分黄金，就如同今天英国女王依然是领土上所有天鹅的主人(99)。假设关于所有权、关于黄金的概念是界定清楚的，这两句话中必有一句表达了事实，但是我们几乎永远无法确定到底哪句话才是真的，不论我们花多少时间进行多么复杂的物理研究。

真的吗？我们可以想象，在某种情况下我们几乎能确定A是真的还是B是真的。如果基于众多有据可查的历史事件，我们发现我牙齿里的金子是有“来历”的，千年以来有人小心翼翼地监控并记录下了它的流向，那么我们可能就会相当明确的认定A是真的。假设我的牙医从博物馆购买了一枚古老的金戒指，那是著名的“凯撒小指之戒”，有众多文献记录可以证明，几个世纪以来那枚戒指从一个国王传到另一个国王直到流转到了那座博物馆，而且还有一盘录像带可以证明牙医融化了这枚戒指并且把金子灌入了我的牙齿石膏模型。这个故事真够古怪的，但它确实是在物理可能性的范围之内。我们还可以设想另外一种情况，假设我是一个淘金狂，旅行到了冰川消融后的阿拉斯加州。这片土地上万年来

都被冰雪所覆盖，我在那里收集了金子并把它们镶到了我的牙齿里。如果这样的话，那么B就变成了真的，而且我们会对此更加确信。但是，如果没有这种既离谱又证据充分的故事的话，那么我们大致应该相信：我们永远也无法知道到底A和B谁是真的。所以无论哪一个是真的，它都是一个无效的历史事实。

还有一个有趣的难题是由量子力学引起的：即便追溯到了我牙齿中所有金原子在这几个世纪中的运行轨迹，只要凯撒戒指上的一颗金原子与来自外界的金原子发生了碰撞，甚至只是非常接近，在原则上我们就无法确定，“碰撞”发生后哪一个原子是来自凯撒的，哪一个原子是来自外界的。与其他事物不同，原子或比原子更小的粒子无法留下指纹或者什么别的可以让我们识别其身份的东西，而且我们也无法持续地追踪它们，所以持续的同一性对于粒子来说并不一定是有意义的，这为我们了解有关那些黄金的事实又增加了一个障碍。

不论整个宇宙是不是决定论的，在设计上，电子计算机却是决定论的，它可以通过数字化而不是模拟来消除亚微观噪声，甚至是随机量子带来的涨落。康威的生命游戏就是一个活生生的例子，不过这种数字化决定论无处不在。数字化是为了决定论的需要，而数字化背后的基本理念是：我们可以有意创造无效的历史事实。把所有主要事件强行划分为两类，比如高和低、开和关、0和1，这让同一类别之间微小的差异被无情的忽略，比如不同数值的高电压、不同特点的开状态，深浅不同的0。这些差异并不决定任何东西，其自身也消失得无影无踪。如此，这些差异对计算机随的一系列状态根本起不到什么不一样的影响。

比如，你的朋友从网上下载了一首歌，刻录到了两张CD中，这两张没有贴标签的CD将用于数字拷贝，我们把一张叫作A，另一张叫作B，但我们当然不会写在上面。你请朋友在漆黑的房间中把其中一张CD里的歌曲拷贝到你的笔记本电脑里。你不要告诉他用的是哪张，并且在你把CD交给他之前要消除掉CD上所有的指纹证据和DNA痕迹。等他拷贝完了，再把两张CD放进一个装满其他CD的袋子里，使劲摇一摇。现在，我们便获得了两个备选的无效历史事实：

A．你的笔记本电脑里的歌是从光盘A拷贝下来的。

B．你的笔记本电脑里的歌是从光盘B拷贝下来的。

对一首歌进行比特流的物理编码时，在微观的精细结构当中总会有

差异。所以当你把一张CD拷贝到计算机内存中时，内存的电压也会具有独一无二的精细结构，如果再把文件从内存拷贝到硬盘或者闪存盘中，它们也会与别的拷贝之间存在微小差异。我们为了方便而称作拷贝的那种过程，其实总会另外制造出一种连续的模拟物理信号，这种信号带有独一无二的精细结构，因为世界上电子和质子就是这么工作的。但是数字化的天才之处就在于，所有这些精细结构都能够在数字化过程中被“忽略”，在“标准修正”中消除。当你拥有了一套标准，就像拥有了一张字母表，每个字母都有多种风格的书写方法，但这不会造成什么影响，因为读起来它们依旧是字母表上的这几个。在电脑中，一切都由0和1书写。

所以，除非那两张CD中有一张在数字化的过程中发生了明显的错误，例如数值写反了，0写成了1或者1写成了0，否则我们无法分辨它们。数字化可以避免将两张CD的个性传播到它的拷贝中，所以经过数模转换，从播放器或者耳麦中传出来的歌声就不再带有CD上的个性。有一些音乐发烧友被誉为“黄金耳”，他们可以听出黑胶唱片和最好的数字唱片之间的差别，可以分辨压缩格式（比如MPEG格式）的音乐文件和非压缩格式的音乐文件之间的区别。但是如果我们要问两段同样的音乐是否出自同一张CD，这些音乐发烧友们恐怕也只能随便猜了。这是一个无效的历史事实，这些拷贝中的细微结构不仅人耳分辨不出，就算用电子显微镜也不行。因为数字化处理已经在物理上杜绝了信息传输的可能，所以要是有人能凭能力分辨不同的话，那他绝对是一位超自然的超感觉能力者。

由于计算机是数字化的装置，所以让它执行几亿步的运算之后再回到与一开始完全相同的数字状态简直是小事一桩，所以，几亿步的数字操作我们也可以一遍一遍又一遍地看它重复执行。

> “等一等，”有人反对道，“你说计算机是决定论的？你说我们能让它反复执行完全相同的上亿步操作？打住吧！要是这样的话，我的计算机为什么这么容易崩溃？为什么我的word周一还用得好好的，周二干同样的活儿时就卡死了呢？”

那是因为你做的不是完全相同的活儿。它卡死了不是因为它是非决定论的，而是因为它周二的状态与周一的状态并不完全相同。这之间你一定干了点什么，要么是打开了一个隐藏“标志”，要么就是激活了word中某个你以前完全没用过的功能。某个发生了改变的数据因此就在程序关闭时被保存了起来，等到周二，这微小的变化就让你的word崩溃了。

要是现在你能想办法把它复位到和周二上午完全一样的状态，你的计算机还是会崩溃的。

那“随机数发生器呢”？我知道我的电脑里内置了一个可以随时制造随机性的装置。

如今，每台计算机都装有内置的“随机数发生器”，它可以根据需要在任何程序中运行。当电脑还没发明出来时，你可以买一本全是随机数字表的书以供研究使用，这一页页的数字全部严格通过了数学家们设计好的随机性测试。当然，对于同一版书而言，每一本中的随机数列都是一模一样的。其中最优秀的一本书是兰德公司于1955年出版的，包含了一百万个随机数字。但是，由所谓的随机数发生器生成的随机数字序列并不具有真正的随机性，它拥有的只是虚假的随机性：它在某种意义上是“数学上可压缩的”，也就是说，这个无限长的数列能够被某个特殊的、有限的机制所把握并计算出来。你的随机数发生器是一个程序，假设它的大小有一兆（800万比特），而它却能够生成一个（每次生成的都是同一个）无限长的数列。如果你想把这个数列发送给别人，不需要在一封无限长的电子邮件里逐个记下每个数字，你只需要把一兆大小的算法程序发给他就行了，通过这个程序就能获得整个数列。以上就是虚假的随机数发生器的基本理念了。

比如，每当你重启计算机时随机数发生器就会开始工作，它永远都会产生完全相同的数列，但是这个数列本身却是完全无规律的，好像它是由真正随机的量子涨落带来的一样。与其说计算机内置了一张“随机数字表”，不如说它装有一盘超长的录像带，录像带中包含了对一个公正的轮盘赌数百万赌局结果的历史纪录。当你开机时，录像带就会倒回到“开头”，这一点有时很重要。计算机程序在不同的“选择”点上利用了随机性，然而，假如初始状态相同，它一遍一遍地重启只能得到同样的状态序列。有时候，如果你想检测一个程序的bug，那么你总是在检测这些状态的同一个“随机样本”，除非你不时地手动在该程序的数据流中插入一个“随机”数。

70

一场计算机国际象棋马拉松

如果决定论是真的，还存在真正的选择吗

要想弄清楚有关决定论与选择的问题是一件非常难的事。如果决定论是真的，还存在真正的选择吗？如果一个看上去具有自由意志的行动者是被决定的，实际上生活在一个决定论的世界中，这一事实会不会抹消他所有的选择和机会？下面给大家介绍一个能帮助我们探讨这一问题的直觉泵，它考察的是一个简化了的人造决定论世界。

假设你在计算机上安装了两个不同的国际象棋软件，然后把它们连到一个监控软件上，让这两个软件进行对奕，一局又一局地下棋，以至无穷。在你关机之前，它们会不会一遍一遍地重复下出同一盘棋局呢？有可能会，如果你那么设置的话。可这样一来你也得不到什么有关A和B这两个程序的有趣的东西了。假设A以同样的方式反复打败B，你也不能由此推断出，一般而言A就是一个更好的程序，或者在别的棋局中A也能打败B。从千篇一律的重复中，你无法得知这两个程序的优点和缺点。你也可以把比赛设置得更复杂些，以便让持续不断的棋局各不相同。这很容易实现。如果在运算过程中，两个程序都借用了随机数发生器，比如，在不存在可知的最优走法，需要"抛硬币"决定的情况下，那么在下一局中，只要你不让随机数发生器在每一局开始时都重新初始化，它给出的值可能就会发生改变，然后程序就会在不同的指令下得到不同的选项，这样一来它就会"选择"不同的棋路。一个变样了的棋局就这样展开了，而第三局也会以不同的方式继续下去。就像不存在两片完全相同的雪花一样，在这一连串对弈中也不存在两盘完全相同的棋局。然而，只要你重启计算机，运行相同的程序，就能得出与上次完全相同的棋局组合，因为决定两个程序如何"抛硬币"的是与上次完全相同的一串虚假的随机数。

那么，假设我们设定好了包括A、B两个程序的国际象棋世界，并研究1 000场比赛，我们会发现很多高度固定的模式。假设我们发现：在1

000场不同的比赛中A总能打败B。遇到这种情况，我们就可能想要一个解释。如果有人只解释说因为程序是决定论的，所以A注定会永远打败B，我们的好奇心一定不会得到满足。我们想知道的是，哪种结构、算法、配置使得A在下国际象棋中具备优势。A显然拥有一些B所没有的素质和能力，我们想知道是哪些有趣的因素在起作用。这也许是一种低层次的解释，比如，可能最后我们发现A和B是两个完全相同的程序，在它们的源代码中有着完全相同的棋步取值器，只不过程序A的编译比B更有效，以至于在同样的计算周期内A可以比B计算的步数更多。也就是说，在下棋时A与B有着完全相同的想法，只不过A计算得更快而已。在正规的国际象棋比赛中都要配备计时器，如果时间到了你还没有移动棋子就会被判输。

但可能我们需要的是一种更高层次的解释：我们平时是怎样讨论如何下棋的？要从这个角度解释A的优越性，比如描述棋局态势，预估可能出现的局面，决定接下来会促成哪种局面，等等。也就是说，A可能会在游戏进行的过程中调整棋子的相对价值，或者它能够更好地评价棋局态势，或者能够随时决定结束某项进程。总之，A与B“想得不一样”，“想得更好、更复杂”。当然，它并不是一个有意识的人，只是这个过程与“想”近似。

实际上，从一个不会总赢的程序里，我们得到的东西可能会更多。假如，A几乎总会赢，并且在它评估哪步棋更好时会使用一套不同的原则。那么我们就需要解释一些更有意思的东西了。要想知道这是什么原因造成的，我们可能要把那1 000盘不同的棋局都研究一遍，从中找寻更深层的模式，我们肯定能发现不少。有些模式对于下国际象棋来说是具有普遍性的，比如，当车老是呆在后方不出时，B几乎肯定会输。而有些模式则独属于A或B，比如，B总爱过早地出皇后。从中我们会发现一些国际象棋策略的标准模式，比如，在同样的局面中，所剩的时间越少，B能搜索到的博弈树节点就越少。简而言之，我们将会发现一块蕴含着无数解释性规律的宝藏，有些是普遍性的（在这1 000局比赛中）、有些是概率性的。

在这场决定论的比赛中，最吸引眼球的就是这些我们可以识别出的下棋模式，但从微观原因的角度看它们几乎都是一样的。从我们的角度看，两个国际象棋程序正在进行紧张的对局。而从“显微镜”的角度，即通过计算机CPU指令和数据流的角度看，实际上只有一台决定论的自动机在以唯一的方式运行着，只要精确地审查随机数发生器的状态和程序

中的其他数据，我们就能预测到它们所下的每一步棋。在它未来的运作中并不存在“真正的”分叉或支流；A和B所做的一切选择都是由那台计算机及其内存的状态所决定的。似乎，这个世界上除了实际发生的事情，别的什么都不可能。假如A在t时刻即将陷入一个将杀网（将杀网一定可以将死对方，但可能不容易被发现），但这时B由于时间不够而停止了搜索，最终错过了关键的一步棋导致将杀网没能建立起来。那么这就意味着，这个将杀网绝对不会发生。如果有人怀疑这一点，我们可以证明给他看：再花一天时间将这次一系列比赛重演一遍就是了。我们将会看到，在同样的时刻和完全相同的局面下，B会再一次由于时间不够而停止搜索。

那么，我们应该怎么说呢？在这个游戏世界中真的没有截与让、没有攻与守、没有错过的机会、没有两个行动者之间真实的拉锯、没有真实的可能性了吗？我承认，像昆虫和鱼一样，我们的国际象棋软件太简单了，它无法承担那种有重要道德意义的自由意志。但是，它们所处世界的决定论状态并不能阻止它们以不同的力量和能力去把握机会。如果想理解在那个世界中发生了什么，我们可以谈谈，实际上是必须谈谈它们的知情选择如何导致局势的改变，谈谈它们能做什么、不能做什么。如果我们想要找到一个因果规律来解释从那1 000盘棋局中发现的模式，那么我们必须严肃地从这一角度出发进行描述：在那个世界中存在A与B两个行动者，它们试图在国际象棋比赛中击败对方。

假如我们再把比赛程序改进一番，让它可以在A获胜时响起一阵铃声，在B获胜时发出嗡嗡声。我们让这场国际象棋马拉松开始，而一个对这两个程序一无所知的观察者可能会注意到，她听到的总是铃声，很少听到嗡嗡声。怎么解释这一现象呢，她想知道。如果不采取意向立场，我们也可以看到并描述A是如何有规律地击败B的，可是即便如此，这种规律性仍然需要被解释。唯一正确的解释或许是：对于“对方在某种局面下将会怎么走”这个问题，A产生的“看法”比B的更好。就此而言，我们要想获得一个解释，就必须接受意向立场。第33章与42章中也有其他一些例子可以说明：如果不采取意向立场，某些因果链条根本无法获得解释。

到此为止说的都很好。但是这里的“决定”和“选择”只是和决定、选择近似的东西，似乎它们少了点什么。真正的选择意味着“可以以其他的方式行动”。但是表象是会骗人的，让我们更细致地考察一个特殊的例子吧。我们可以在程序比赛中加入第三个下棋程序C，C比A和B都要厉

害，几乎每次都能打败它们。让我们假设，有两局比赛前12步棋是完全相同的，后面的则不同，在这两局中C分别击败了A和B。专家们聚到一起回看了比赛，他们一致认定：在第12步棋中，也就是两局比赛共有的最后一步棋，如果A或B选择了王车易位，那么它们很有可能会打败C。第12步的王车易位是获胜的关键，而A和B都把它错过了。

程序A的设计者耸了耸肩说：“其实，A本来可以王车易位的。”B的设计者插话说：“我的也是，B本来也可以这么下的。”但A的设计者说的是对的，而B的设计者说的是错的！这怎么可能呢？整个比赛程序T都是确定性的，如果我们再来一遍，重新回到与上次比赛完全相同的那个状态，无论A和B都仍然不会做出王车易位的选择。那么，程序A的设计者是不是在骗自己呢？这可不一定。当我们问A是不是本来能够王车易位时，我们需要考察什么？一遍一遍地看与此情况一模一样的棋局对我们一点帮助也没有，我们应该查看那些与此情况有相似之处的棋局。一旦我们发现，当在其他比赛中碰上类似情况时，A确实会算得更多一点，它能够发现王车易位的好处并下出这样一步棋，那么“A本来可以王车易位”这个说法就得到了支持。

我们或许会发现，随机数发生器中仅仅一比特的变化可能就会导致A的王车易位。假设A的设计者再往深处挖掘一下，把A过早停止“思考”（因为时间有限，不论多么聪明的国际象棋程序，也不得不在某个点上强行停止它的搜索）时的实际运行状态揭示出来：A想到了王车易位，并分析走这步棋的后果，可是由于时间有限，它让随机数发生器帮它掷硬币决定自己该不该继续分析，实际上，到此为止所有的步骤都是最佳的，但是掷出的结果是不该。假如掷出来的那个随机数是1而不是0的话，A就会继续再思考一会儿，它就会进行王车易位了。“只要这个随机数变一下，A就赢了！”A的设计者说到。如果是上面所说的这种情况的话，我们可以说，A没有进行王车易位是一个意外，因为随机数发生器没有带来好运。

下面我们看看B，它的设计者也认为“在那个状况下B本来可以进行王车易位”，但却没有证据能支持这一说法。的确，B“知道”在哪种情况下进行王车易位是符合规则的，可能它也曾“动过念头”，但是这离真的做出王车易位还远着呢。在那个局面下，王车易位是一步很深奥的棋，报纸的国际象棋专栏会往这种走棋后边加上一个“！”，它已经远远超出了B的分析能力。现在，我们拥有了一个完全被决定的世界，即程序T，在这个世界中，A本来可以进行王车易位，而B则不能。A与B之间的差

别是真实的，是可以解释的，它们有着能力上的不同。我们可以用一种明显自相矛盾的方式把这一点表示出来：

> 在t时刻，A本来可以进行王车易位，但是，在时刻t，宇宙中却不可能有王车易位的事件发生。

这样描述的意义是什么？很简单：只要我们把A和与A紧密联系的周边环境分开来看，也就是说，不考虑随机数发生器，那么无论A是否进行王车易位就都不是被决定的了。严格来说，A是否被决定依赖于A的外部环境。基于给定了的外部世界在t时刻的运行方式，A不可能进行王车易位，但这“不是A的错”。B与之相反，它本来就无法进行王车易位，它在本质上走不了这步棋。要想让B走出这步棋，我们需要改变太多的事实。

这一发现对我们很有帮助，即并不需要在非决定论的状况下，A和B“本来能做的事”之间才有区别。即便在一个决定论的世界中，我们也能看出A能做某些B做不了的事，我们解释“A为什么会打败B”时也要解释这种差别。事实上，因为决定论在这个游戏世界里是正确的，A和B在特定情况下都只能做它们实际能做出事，只要完全相同的情况再现，它们还会一遍又一遍地重复，所以，如果我们要对那个显而易见又非常客观的规律“A打败B”做出解释的话，上述事实既没意思且无关紧要。

国际象棋软件不是道德能动体，它所做出的选择不承担任何道德责任，因为它的世界完全是非道德的。对一个国际象棋软件来说，违反国际象棋规则是不可想象的，因此也就不存在对违反规则的惩罚。但是就像我们刚刚看到的，即便在一个简单的、决定论的电脑国际象棋世界中我们也能对A和B做出真实而有意义的区分。当A做了件傻事或聪明事时，我们可能会说：“A本来可以不这样做，但B不行。”如果你认为我们说错了，认为“因为那个世界是被决定的，A与B永远都不能不这样做”，那么你就错了。

就像我们刚刚看到的，A与B在下棋能力上有差别，而“本来可以不这样做”捕捉到了这种差别的一个侧面。那么如果事关道德又会怎样呢？有些人做了坏事，人们会说“他们本来可以不这样做”，并以此为由不宽恕他们；但他们也同意，即便换了别人，在类似情况下也可能这样做。他们如此认为并没什么错，因为这与决定论是否正确没有关系。这里展现出的是一个道德能力上的真实差异，人们的反馈也会因这种差异而不同，它不取决于决定论或是非决定论。

为了更清楚地说明这一点，让我们站在程序B的设计者角度思考一下。她想要知道自己是否已经找到了B的弱点所在。在另一场比赛中，由于没有进行王车易位，B又输了；那么，这次的B是否本来可以王车易位呢？如果我们发现这次的失败全是拜随机数发生器中一个字符变化所赐，那么我们也许就不必再做任何设计上的改进。如果在相似的局面下，B往往能够走出王车易位这步棋，那么它也许已经变得和我们所期望的一样好了。程序必然会用到随机数（就像掷硬币）来终止搜索、继续游戏。因此，在发现关键的一步棋之前恰好停止搜索这种事肯定会发生，这全拜掷硬币所赐。请注意，即使我们让程序B或A使用量子随机数发生器，比如一台盖革计数器，它基于亚原子微粒的随机轨迹来生成随机数，情况也不会有什么改善。仔细想一想我们说的情况，B没有王车易位是因为在应该掷0的地方却掷了1。如果随机数发生器生成了一个0，B进行王车易位；如果生成的是1，不王车易位。“B本来可以王车易位的”，当生成的是1时一个观察者会这么说。是的，可所有这些都不是B所能决定的。在出现这种状况的棋局中，B进行王车易位与不进行王车易位的概率是一半一半，不管它使用的是“真实的”还是“虚假的”随机数发生器。

哲学家大卫·威金斯（David Wiggins）曾经写过决定论下的“宇宙不平等”（Wiggins，1973，p. 54），但是我们这个关于计算机国际象棋比赛的直觉泵告诉我们，在非决定论下有着同样的“宇宙不平等”。B任由它的随机数发生器或者伪随机数发生器的“摆布”。当然，A也一样，我们都一样。真随机数发生器没有任何理由得到青睐，当然，除非你准备跟全知的神对弈，而神能够看透你的伪随机数发生器并做好相应的计划！

所以，如果我们还是希望非决定论是真的，就该找别的理由了。或许，我们不需要诉诸非决定论就可以拥有有价值的自由意志。等等，还有一个理由：

> 我不能改变过去，但如果非决定论是真的，我就能够改变未来！

不是的。你想把未来从什么改成什么呢？从将要发生的改成将会发生的？你无法改变过去，也无法改变未来。因为改变未来这个概念是不合逻辑的。所以：

如果决定论是真的，我们不能改变未来；如果决定论是假的，我们也不能改变未来。所以结论就是我们根本没法改变未来。

为什么我们有要去改变未来的想法呢？因为我们希望自己能够预见到灾难，这样我们就能通过努力来避免它的发生。可是，无需非决定论我们也能够做这样的事。如果有人向你扔了一块砖头，你看到并躲开了，那么你就能够避免被那块砖头击中，这很了不起。这块砖头将要砸到你的头上吗？在一个意义上说，是的，因为根据轨道来看很明显这块砖头会正好落在你的脑袋上，但由于你看到了它（由它反射的光进入了你的眼睛，因此你看到了它，你的大脑计算到了危险，因此你肯定会采取行动），所以就把它躲开了。你当然也可以不躲避它，只要你愿意，如果你有理由认为被它砸中对你来说更好的话。或许直到最后一刻，一些观察者都无法知道你会不会被砸中。所以如果他赌你会躲，也有可能输。回到我们前面讲过的那个理由——不想被预测到，但这并不需要世界是非决定论的。

这个直觉泵的目的是什么呢？它从一个有意义的角度让我们看到，与那个普遍的、未经检验的观点不同，人们常说的那句短语“本来可以不这么做”，并不依赖一种非决定论。即便“本来可以不这么做”有一种既不兼容于决定论又具有道德重要性的意义（也就是说，不仅仅为了满足形而上学的好奇心），那这也仍是需要证明的，而且证明的责任在那些相信它的人一边。我们又发现了一个“明摆着”的观点其实根本就不是那么地“明摆着”。

71

终极责任
斯特劳森的错误论证

到目前为止，我们已经考察了一些简单的选择：石头剪子布、下棋、躲石头，它们都与道德责任无关。我们不应该只考察棋类游戏或弹跃的瞪羚这类意向系统，是时候该具体地考察我们想要成为一个道德行动体的愿望了，也许这样我们才会真正需要非决定论。许多思想家都是这么想的。对于他们而言，那些思想实验只是在分散我们的注意力。有些思想家会拿下面这则清晰的论证来对阵决定论观点，哲学家盖伦·斯特劳森（Galen Strauson）把它整理成了如下形式（Strawaon, 2010）：

① 在任何给定情况下，你因自己是什么样的人而做出你所做之事。

② 因此，为了在根本上对你所做的事负责，你必须在根本上对你是什么样的人负责，至少对你的心智的一些关键方面负责。

③ 但是在任何方面，你都无法在根本上对你是什么样的人负责。

④ 所以，你无法在根本上对你所做的事负责。

第一个前提是无可置疑的：“你是什么样的人”包括了你在时间中的全部状态，无论你是如何进入这些状态的。不论这些状态是什么，不出意料的话，你的行为都得从它们当中产生。第二个前提论述的是，你无法在根本上对你所做的事负责，除非你能在根本上（至少能在某些方面）对你所处的状态负责。而第三个前提告诉我们这根本不可能。

所以，第四个结论看上去似乎符合逻辑。不少思想家认为这个论证很严格并且非常有意义。真的是这样吗？让我们来仔细考察一下第三个前提吧。为什么我们不能（至少在某些方面）在根本上对自己是什么样的人负责？在日常生活中我们会明确区分要不要负责，这点在道德上也非常重要。假如，你设计制造了一个机器人，然后不加看管地把它扔到

外面的世界，因为你完全清楚它所有可能会参与的事情，并且估计它可能会把一个人弄成重伤。你需要为这起事故承担责任吗？大多数人都会说，是的。你制造了它，你应该预见到它的危险，实际上你也确实预见到了，所以伤害发生后，被责备的人就是你，你至少要承担一部分责任。如果你坚持认为自己对机器人造成的伤害根本不用负责，则很少有人会对你表示同情。

现在，请思考一个稍有不同的例子：你设计制造了一个人，比如晚些时候的你自己，然后把自己扔到了充满危险的世界，你完全清楚自己可能会遇到的危险。那个你在酒吧里把自己灌醉，然后开着车上路了。当那个你撞上一辆校车时，难道你一点儿也不用对“这样的你”承担一部分责任吗？常识告诉我们，当然要承担。（那个酒保或纵容你的酒吧老板或许也要承担一些责任。）但是，在斯特劳森强有力的论证面前，这怎么可能啊？很简单，请记住，斯特劳森说的是我们不能对自己是什么样的人负有绝对的责任。可是这又怎么样呢？谁会认为我们必须承担绝对的责任？这简直是一种根本不可能存在的状态，即便非决定论是真的也不可能！根据这条论证让我们期待非决定论也站不住脚。斯特劳森是这么说的：

> 要做到对所做之事绝对负责，人必须是自因，自己是自己的原因，但这是不可能的。即便我们不完全是由物质组成，即便我们有不朽的灵魂，这也还是不可能的。（Strawson, 2010）

绝对责任是干扰视线的伪焦点，是没人渴望得到的祝福。斯特劳森不这么认为，他批评我忽视了下面这一点：

> 他没有建立那种大多数人都想去相信、或者真的相信的绝对的自由意志和道德责任。他做不到，他自己也知道这一点。（Strawson, 2003）

他说的完全正确：我没有建立那种大多数人都想要相信的自由意志，我自己也知道。不过，我认为大多数人的那种想法是错的，如果他们真的相信那也相信错了。论证的责任在斯特劳森等人一边，他们得证明为什么我们应该在生活中关心根本的责任，或者关心决定论/非决定论的问题。他们可以定义各种和决定论不相容的概念，展示出人们觉得

它如何重要，但是他们也必须说明这些人并没有自我欺骗。为什么有人会关心这一点？请注意我是在反问。我已经做好被打脸的准备了。我很想看到斯特劳斯或者谁站出来回答这个问题，但到目前为止我们还没有这样的志愿者。

在结束对斯特劳森论证的讨论之前，我想问大家一个问题：你们是否注意到，他的论证与我们之前看到过的一个论证惊人地相似。下面我仿照斯特劳森的形式重写一下那个论证，让大家看到它们的相似之处。

① 在任何给定的情况下，哺乳动物因为是哺乳动物而是哺乳动物。

② 要想是一个哺乳动物，你的母亲也必须是哺乳动物。

③ 但是，你的母亲也必须符合这一条件，还有你母亲的母亲，以至无穷，而这是不可能的。

④ 所以你不可能是一个哺乳动物，因为哺乳动物在根本上是不可能的。

你应该永远警惕这种“祖先”论证。几乎可以肯定，这是一种经过伪装的，被称作连锁推理（或者“谷堆”论证）的古代谬论：

① 一粒小麦不是谷堆。

② 在一粒小麦旁加上一粒小麦也不是谷堆。

③ 你不能通过增加一粒小麦，而使非谷堆变成谷堆。

④ 因此，不存在谷堆这种东西！

几千年来，哲学家们对连锁推理悖论以及其中所用术语的界限模糊问题（该悖论俨然依赖这些界限模糊的术语）著述颇多，但至今我们仍然不知道该如何判定和避免这一谬论。（你们可以登陆斯坦福哲学百科，看看他们最近做的一个出色的调查问卷。）甚至还有些勇敢的哲学家声称连锁推理是有效的，他们试图接受“既不存在秃顶的男人又不存在不秃顶的男人”这样的“事实”。这个立场捍卫起来好难！但就像43章所讲的，达尔文已然教会了我们如何拒绝连锁推理；我们无需在我们的族谱中画下一条原则上的物种分界线。

据我所知，我是第一个指出斯特劳森的论证（以及其他有关的文本中出现的相似论证）与那个古代谬误有相通之处的人，可事实就是如

此。我想，就像爬行动物、兽孔目动物经过世世代代的进化可以变成哺乳动物一样，人类也是从婴儿到成人一路走来，渐渐成长为一个可以承担道德责任的人。你是一个哺乳动物，不必是一个绝对的哺乳动物；你可以承担责任而不必承担绝对的责任；你也可以拥有一种值得向往的自由意志，而不是永远绝对的自由意志。实际上，绝对的自由意识就像奇迹一样，你必须给出一个强有力的论证才能证明为何人们要贪图这样的东西。难道你想成为上帝吗？真糟糕，我们一般可没有这种好运，不过，能成为第二好的也不错呀。

72

掘土蜂状

有限能力无法成就自由选择

侯世达（Hofstadter， 1982）杜撰了“掘土蜂状”（sphexishness）这一术语，指容易被人误以为是大智若愚的一种死板的、机械的无意识状态。最典型的例子，也是这个词的来源是一种有着古怪行为的黄蜂。侯世达和我都被迪恩·伍尔德里奇（Dean Wooldridge）的《大脑机制》（*The Machinery of the Brain*）这本科普书中的一篇文章惊到了。文中这样描写掘土蜂：

> 产卵的时节，掘土蜂会为后代挖掘地洞，然后寻找并蛰咬一只蟋蟀，它的蛰咬会令蟋蟀麻痹。接着，它会把蟋蟀拖进地洞，放在卵的周围，再把洞口封死，然后头也不回地飞走。在恰当的时候，蜂卵孵化出来，掘土蜂的幼虫会以被麻痹的蟋蟀为食，此时蟋蟀还没有腐烂，这个地洞就像是掘土蜂的冷藏柜。以我们人类的眼光来看，这套精心组织、似乎目的明确的过程传达出了一点逻辑的味道，甚至还有些思想性，直到我们发现了更多的细节。比如，掘土蜂会例行公事地把麻痹的蟋蟀拖到洞口放下，然后进入洞中检查一番，确认一切都好之后再从洞里出来，最后把蟋蟀拖进洞中。如果当掘土蜂进入洞中进行初步检查时蟋蟀移动了几厘米，它出来后会再把蟋蟀拖回到洞口而不是洞内，然后把“准备工作”重做一遍。掘土蜂从来也没想过可以直接把蟋蟀拖进洞里。有一次这一过程重复了40多次，每次都是同样的结果。（Wooldridge, 1962, p. 82）

不管怎样，我们最近获悉，就像所有科普作家常做的那样，伍尔德里奇对这一现象作了一番过于简化的概述。心理学家拉斯·奇卡（Lars Chittka）在给我的一份报告中引用了法国自然学家《昆虫记》作者让-亨利·法布尔（Jean-Henri Fabre）的一些研究成果，那明显是伍尔德里

奇观点的出处。如果伍尔德里奇读过法布尔的书就会知道，实际上只有某些掘土蜂是“掘土蜂状”的。而且，法布尔实际上也很想指明这一点：让你乍一看以为掘土蜂很聪明，再一看以为掘土蜂很笨，最后一看又发现有些掘土蜂一点也不那么“掘土蜂状”。奇卡给我的是法布尔《昆虫记》的德语版本，里面是这样说的：“两三次之后，掘土蜂用钳子揪着猎物的触须，把猎物拖进洞里。现在谁是傻瓜呢？”

所以，“掘土蜂状”其实有点用词不当，不过既然都这么叫了，我就不得不继续让掘土蜂忍辱负重。从某种意义上来说，它们很幸运，可以以这种方式成为人们关注的焦点，也算是很好地为自己提高了适应性吧。想想看，你会投票保护谁的栖息地，掘土蜂还是humdrumbeetle？也许掘土蜂没有大象、老虎和狼那样“有魅力”，但因为这成问题的“掘土蜂状”也使得掘土蜂名声大噪。

“掘土蜂状”是一种很重要的性质，与其说是因为所有例如昆虫、蠕虫、鱼等简单动物都在不同程度上具有这种性质，倒不如说是因为它让我们找到了一个术语，用来指代一种有限的、机械的、短视的能力，而更奇妙、更复杂、更有理解力的心智则建于其上。任何心智的构建模块都最好是掘土蜂状！或者，就像我之前提到过的，构建模块应该是近似心智，应该是我们心智中苍白的影子。我们也可以用掘土蜂状把有道德能力的心智与没有道德能力的心智区分出来：如果一个人的心智到了像掘土蜂似的程度，不管这是因为脑瘤或脑损伤，还是因为严重的神经失调或是精神病，或者纯粹是因为愚昧或不成熟，那么他也就无法做出有意义的事。

为了揭示出“掘土蜂状”这个性质，执着的生物学家们不惜一遍又一遍地搅扰掘土蜂，他们就是我们着实该担心的那种典型的操纵者。在哲学上，许多关于自由意志的思想实验都引入了这样一种傀儡操纵者或者一个毒恶的神经外科医生，他们偷偷地让某个人遵循他们的指令。这些可怕故事的寓意大概是：即使实际上并不存在这种傀儡操纵者，但各种各样的环境因素经由我们大脑的处理将会决定我们的行为这一事实，也证明了这种傀儡操纵者是可能存在的。萨姆·哈里斯（Sam Harris）的那本《自由意志》（*Free will*）的封面上就画着一组操纵傀儡的线。[100]但这显然是一种不当的推论。如果环境借助我们运转良好的感知系统和未受欺骗的大脑“控制”了人，那也没有什么好怕的。相反，倒是没有比这更令人满意的了，我们周围的事物和事件导致我们产生有关它们的真信念，借助这些信念把我们的行为调整得对自己更有利。光子从滩涂的气

孔上反射到我眼中，让我拿起蛤耙和篮子开始挖蛤蜊。如果这算被环境所控制的话，那我举双手赞成。像大多数人一样，如果我的朋友请我吃丰盛的美食，我并不觉得自己受到威胁或被操纵，因为我知道我无法抗拒美食的诱惑。

在那些以傀儡操纵者或恶毒神经外科医生为例论述的故事当中还有另外一件值得我们注意的事，你会发现这种操纵永远永远是秘密进行的。为什么会这样呢？因为只有当我们的行为或选择不知不觉地被那些神秘的操纵者所左右时，我们才会在直觉上感到我们不具有自由意志。其中的道理并不难理解，让我们回想一下开启了博弈论的那个洞见：当一个行动者意识到另一个行动者想要操纵他时，他会寻找对策，或者至少会去调整他的行为以应对这个情况。两个行动者之间的竞争性互动涉及许多层面的反馈，因此会削弱想要成为操纵者的人的控制力。而一旦操纵没能秘密进行，而且还正中“傀儡”的下怀，那么局面就完全反转了。

我们只需对标准的思想实验稍加改动，转动直觉泵上的旋钮，就能演示出上面那种情况。哲学家哈里·法兰克福（Harry Frankfurt）发明过一个直觉泵，讲的是一个神经外科医生偷偷把一个可以监控神经刺激的设备植入了一个人的大脑，以确保那个人所做的一切决定都在医生的掌控之下（Frankfurt, 1969）。假如你就是那个人，如果你在一个两难的局面中选择了A而不是B，且那个医生就想让你这样选，那么他会袖手旁观；但是如果监控装置显示你想要选择B，那么外科医生就会按下按钮阻止你选B，所以无论如何你都会选择A。你对这一切都没有感觉。在这两种情况中你的选择都不是自由的吗？多年以来，哲学家们为了解决“法兰克福案例”和其变种写过成百上千的论文。其实，只需转动一下旋钮，局面就会产生巨大变化，据我所知这还从未有人冒险尝试过。下面请大家过目：

为了抑制你对甜食无尽的爱，你和一位医生签了一份合同，他帮你植入一个监控装置，并且你还会支付优厚的报酬让他监控你的每一餐。这为你提供了一个安全网以防止你点餐时选择热巧克力圣代和芝士蛋糕。你们俩都希望他永远也不必按下手中的按钮。不久之后，你都快忘记了自己还被他或他的助手控制着。你已经被“治”好了，这个小装置已经连续记录下了成百上千个指令，显示着你选择了A，“不了，只需要一杯黑咖啡，谢谢”。这能不能作为你自由选择的例子呢？为什么不能？对你来说，让自己多少有些脆弱的意志力得到点儿帮助难道不是一件很明

智的事吗？

73

巴西来的男孩
一个干扰人们思考的“坏”直觉泵

心理学家约书亚·格林（Joshua Greene）和哲学家乔纳森·科恩（Jonathan Cohen）在2004年合写了一篇题为《对法律而言，神经科学既改变了一切又什么也没改变》（*For the Law, Neuroscience Changes Everything and Nothing*）的文章，发表在了著名的《皇家学会哲学汇刊》（*Philosophical Transactions of the Royal Society*）当中。这篇文章唤起了一场由科学发现引发的法律革命。

> 法律声称自己的前提不外乎是一种形而上学的适度的自由意志观念，它可以与决定论完美相容。但是我们认为，法律的直觉支撑完全是建立在一种过于形而上学的野心之上的，因为自由意志主义者的自由意志观念正遭受着决定论的威胁，更直言不讳地说，它正遭受着即将到来的认知神经科学的威胁。（Greene & Cohen, 2004, p. 1776）

他们把情况说得很微妙，这是必然的，因为他们既认为自己有丰富的论据来支持相容论（Compatibilism），这也是我在这本书里所捍卫的观点，但同时他们也想要表明一下，关于自由意志实际上“我们都还举旗不定”。他们提出并设计了一个思想实验，旨在揭示我们日常生活中的常识确实依赖着自由意志论。请注意，这个思想实验很像我前面要求过的那种论证：如果您认为我们应该关心自由意志论，请做出说明并给出理由。他们的思想实验是受一部电影《巴西来的男孩》（*The Boys from Brazil*）的启发，影片讲述的是一个培养希特勒克隆人（多亏抢救出了一些DNA）的纳粹科学家。下面请看这个思想实验：

> 让我们假设，一群科学家成功地创造出了一个人，就叫他傀儡先生好了，他们的设计令傀儡先生做出了犯罪行为，比如

在毒品交易失败时杀人。（Greene & Cohen, 2004, p. 1780）

关于这个思想实验，他们说：

> 是的，他在犯罪时与其他罪犯一样是清醒的，是的，正是他的欲望与信念导致了他的行为。但是，这些信念和欲望是受外力操纵的，而这也是为什么在直觉上我们觉得他更值得同情而不是谴责……那么，在我们和傀儡先生之间到底有什么区别呢？一个明显的区别是，我们认为傀儡先生是一场邪恶阴谋的受害者，而大多数人却不是。但这重要吗？认为傀儡先生不该承担全部责任的理由是他的行为由外部原因所决定……事实上，这种外部力量与邪恶科学家的欲望和意图有关，但这并不重要，不是吗？重要的是，这些力量超出了傀儡先生的控制，它们并不是傀儡先生的。（Greene & Cohen, 2004, p. 1780）

你们怎么想？对他们论证的目标而言，这是一个好的直觉泵吗？或者不是？有意思的是，作者意识到了这个值得思考的问题，他们写到："丹尼尔·丹尼特可能会反对说，傀儡先生的故事只是一个有误导性的'直觉泵'。"我确实是这么想的。我会把它归为一种吊杆托架。可是，他们耸耸肩继续说道："在我们看来，人们对傀儡先生和他的生活了解得越多，就越不会认为他真的该对他的行为负责，越不会认为他应当因此受罚。"

来，让我们仔细看看，转动直觉泵上的旋钮，实际上都是什么在运作。下面，我准备转动四个旋钮。

第一个按钮：让我们拿掉邪恶的阴谋，作者自己都坚持认为它不重要。我很奇怪他们为什么会冒失地在故事中加入"邪恶的神经外科医生"，并设定他们显然是无罪的，这种做法总是令我起疑。所以现在，我要用"冷漠的环境"来代替"一群科学家"，看看他们的结论还是否正确：

> 转动前：让我们假设，一群科学家成功地创造出了一个人，就叫他傀儡先生好了，他们的设计令傀儡先生做出犯罪行为，比如，在毒品交易失败时杀人。
>
> 转动后：让我们假设，冷漠的环境成功地创造出了一个人，就叫他傀儡先生好了，他们的设计令傀儡先生做出犯罪行为，比如，在毒品交易失败时杀人。

第二个旋钮：故事里已经没有阴谋了，所以我们不得不把“他们的设计令”也去掉，添上“有很高的概率”：

转动前：让我们假设，冷漠的环境成功地创造出了一个人，就叫他傀儡先生好了，他们的设计令傀儡先生做出犯罪行为，比如，在毒品交易失败时杀人。

转动后：让我们假设，冷漠的环境成功地创造出了一个人，就叫他傀儡先生好了，他有很高的概率做出犯罪行为，比如，在毒品交易失败时杀人。

第三个旋钮：我想改变犯罪动机，仍然是杀人，但却发生在一个相当不同的场景下（这应该不重要，对吗？）。

转动前：让我们假设，冷漠的环境成功地创造出了一个人，就叫他傀儡先生好了，他有很高的概率做出犯罪行为，比如，在毒品交易失败时杀人。

转动后：让我们假设，冷漠的环境成功地创造出了一个人，就叫他傀儡先生好了，他有很高的概率做出犯罪行为，比如，为了掩盖挪用公款而杀人。

第四个旋钮：让我们改动一下罪犯的名字，毕竟只是改个名字而已。

转动前：让我们假设，冷漠的环境成功地创造出了一个人，就叫他傀儡先生好了，他有很高的概率做出犯罪行为，比如，为了掩盖挪用公款而杀人。

转动后：让我们假设，冷漠的环境成功地创造出了一个人，就叫他自主队长好了，他有很高的概率做出犯罪行为，比如，为了掩盖挪用公款而杀人。

现在它激发出的是什么直觉？还是同样的直觉吗？你还觉得他更值得同情而不是谴责吗？如果细节再充实一点，可能我们就会看得很清楚了。下面就是我友好的修订案，请格林和科恩过目：

自主队长在哈佛大学主修经济学，毕业之后去了雷曼兄弟

公司上班，和他一起工作的同事都在偷奸耍滑并以此牟利。他爱上了一个傍大款的迷人女郎，她威胁说，除非自主队长马上变得很有钱，否则她就甩了自主队长。自主队长看到有个机会可以挪用公款而不被发现，基本上可以做到隐形犯罪，于是他决心冒险一试。唉，可惜人算不如天算，冒出来了一位目击者，那人错就错在不该站得离顶楼花园的栏杆太近了……一次快速的“跌倒”——啊！——证人摔死在了大道上。他引起了人们的怀疑，不久之后自主队长就被捕了。

到这里，你还倾向于认为，由于他的行为是由外部原因决定的，所以他“其实不该负责”吗？即使你仍然愿意把自主队长看作是奢靡环境的受害者，我想你也不得不承认，与之前所描述的那种力量相比，他所受诱惑的力量已经大大削弱了，甚至只能算是前者那种力量的某种余波。或许我其实只是在利用你们的感情让你们支持相反的论点。我并不是想通过这套改变来证明，尽管人们是被决定的，他们也要或者也能对自己的行为负责，我只是想说明格林和科恩给出的那个直觉泵一点也不可信，因为其中可用的、允许转动的旋钮设置大大地干扰了我们的判断。也许设计它的初衷并不是为了糊弄人，但它确实妨碍了我们清晰的思考。

第七部分小结

人们非常在意自己拥有的自由意志，不过对于什么是自由意志或它本该是什么，人们似乎也常会被一些观念所误导，就像他们在颜色和意识的问题上被误导一样。我们的决定并不是发生在大脑中的小小奇迹，它也并不违反那些可以解释我们身体其余部分的物理和化学原理，即使大多数民众都以为“如果我们的决定真的是自由的，它就必须是个奇迹”。但是，我们也不能因此推断我们没有自由意志，因为在这个疯狂的意义上的自由意志并不是自由意志的唯一概念。在法律中，人们会区别“出于”自愿签订的合同与在强制、幻觉或精神错乱之下签订的合同。我们非常熟悉这个意义上的自由意志，它蕴含在构成了我们常识映像的实践与态度之中，没有任何迹象显示它的各个特点是依附于那个疯狂的意义上的概念。

几百年来，一直有哲学家坚持法律意义上的这种自由意志才是重要的、是我们应该关心的，而它相容于决定论和唯物主义、相容于物理学和化学无可争议的统治。这一部分所介绍的直觉泵以及其他思考工具旨在支持和推进对相容论的理解。多年以来，人们提出过各种版本的相容论，相容论不仅是哲学家，还是法官、律师和其他一些人之间的共识，他们认为在“谁应当负责”和“谁应当受到宽恕”之间的区分，取决于人们在行动时是否具有自由意志。现如今，一些科学家正在挑战这一共识，他们当然也有权利这么做，就让我们严格地审视一下这些论证吧。

或许，科学正在让我们意识到一些关键的，甚至是具有革命性的东西：人们永远不该对他们所做的一切负责，对值得赞许的行为与值得谴责的行为之间的区分没有稳固的基础。但是，这样一个革命性的结论更需要人们认真地对待每一个细节，可迄今为止科学家给我们看到的那些论证还远远不够。恶毒的神经外科医生把她的病人变成了残废，凭的不是别的而是一个虚假的观念；有影响力的科学家们颁布的错误观点可能会剥夺自由意志的多样性，它们原本合情合理并且对我们的生活有益。我呼吁，每一方都应当小心谨慎。

相容论虽然在哲学家当中非常流行，但也常常会引起质疑。哲学家伊曼努尔·康德（Immanuel　Kant）曾把相容论称为一个“可怜的托词”，

而今天的作者们也常常对我们这些持有相容论的人是否真诚表示怀疑。事实上，他们也理当如此。科学教导我们，要特别警惕那些一厢情愿的想法，我们在科学研究中设定的许多规则就是为了防止我们被自己的希望所骗时还以为是被证据所说服。让我们想象一下，假如有些天文学家宣称一颗巨大的小行星将在10年后撞击地球，消灭地球上所有的生命，而后又有一些天文学家宣布他们再次对数据进行了分析，让我们可以松一口气，因为那颗小行星会跟地球擦肩而过。我们怎么才能知道第二批天文学家是不是在自我欺骗，又或者只是告诉了我们一个善意的谎言？要对他们的计算进行一遍一遍的验算或自己重新再做一番独立的计算，不要因为他们的结论没有明显的错误又能令你满意就美滋滋地接受它。但是永远不要忘记，他们也有可能是对的。不要矫枉过正，不要“在原则上”怀疑所有看起来“好得令人难以置信”的事。

相容论真是好得让人不能信吗？那倒不是。我认为它是真的，而且我们既可以完全彻底地摒弃那些杞人忧天的想法，同时又改变和修正我们的理解，理解是什么支撑着我们的道德责任观。但这也是一项我们在未来才能完成的任务，它还需要很多人的努力。对我而言，它是我们今天面临的最困难也是最重要的哲学问题。因为赌注很高，问题很棘手，我们的判断也容易被情绪所扰。我们需要调动所有的思考工具，可这还不够，在前进的道路上我们还必须制造更多。

第八部分

做个哲学家是一种怎样的体验

INTUITION PUMPS
AND OTHER
TOOLS FOR THINKING

每当我们离一件东西太近，就会很难看清它的面貌。近年来，我和其他哲学家的交流比以前少了，我花了更多的时间来与科学家或者别的思想家交流。但我仍然是一位哲学家（先不管别的哲学家怎么说我！），而且我也一直愿意接受挑战，向非哲学家解释为什么哲学值得研究。对于一些完全外行的人来说，哲学常常是可笑的，简直是“聪明劲没处使”的典型。所以他们错过了一些东西。想一想史特金

定律：任何事物中的百分之九十都是垃圾，而他们却还在错过那剩下的百分之十。

入哲学一行50余年了，我熟悉它，现在，我把更多的时间花在了哲学之外，也让我清楚地看到了它令我费解的一面。科学界的一些朋友和同事向我坦白，说他们一直不明白我为什么不跳槽，不加入他们的行列。简单的回答是：我想利用这种跨界把两边的好处都占上。与科学家在一起工作，给了我很多有趣又棘手的事实以供思考；而保持我的哲学家身份，则令我即使没有实验室和研究经费也能钻研一切理论和实验，甚至连培养皿都不用刷。引用我最喜欢的一首乔治·格什温的歌吧："那是一份好差事，假如你能得到它。"

科学家们近些年来比以往更尊重和关注哲学了，在我看来这是一个好消息，在我主要研究的心灵哲学领域尤其如此。认知科学家所研究的几乎就是哲学家们几个世纪以来一直在思考的问题：感知、记忆、意义、意志和意识。而有些哲学家也在科学界获得了关注，因为他们主动了解了相关的科学进展，为澄清和推进科学研究提供了有益的建议，同时也更好地向科学界以外的听众解释了科学研究的结论。但是当这两群人试图沟通时，交流上的失败却大量存在。为了今后我们能够更好地相互理解，我打算在本书的这一部分打消一些双方的分歧。

74

浮士德式交易

哲学家的目标不同于科学家的目标

以前我总爱给我的哲学家同行们提出这样一个问题：如果靡菲斯特（Mephistopheles）[(101)]给你以下两个选项，你会选择哪一个？

A．你彻底解决了一个重要的哲学问题，经过你的努力这个问题大家以后就不用再讨论了。多亏了你，哲学的某个领域彻底终结了，你因此在哲学史中占据了一定地位。

B．你写出了一部巨著，它令人感到既着迷又困惑、充满争议，在接下来的几个世纪中这本书都在学者们的必读书目之列。

一些哲学家不情愿地承认他们会选择（B）。如果必须二者择一，他们宁愿选择被阅读而不选择正确。就像作曲家、诗人、小说家以及其他艺术创作者一样，哲学家也希望自己的作品被人阅读，被上百万人（如果有可能，数十亿人更好！）不断地阅读。但他们也得往科学家们的要求上靠拢，毕竟，哲学家还是应该试着把握真理的。

而当我向科学家们提出同样这两个问题时，他们却常常毫不犹豫地选择（A），对他们来说这是明摆着的事嘛。当他们听说这个选择在很多哲学家看来难以抉择，甚至有些人会蠢到选（B）时，他们会惊愕得直摇头。但科学家们的这种反应，却忽略了一个重要的事实。尼古拉斯·汉弗莱说过：

> 在《两种文化》（*The Two Cultures*）一书中，斯诺（C.P. Snow）用“科学中的莎士比亚”来赞美科学中的伟大发现，但在某种意义上，他却犯了一个根本的错误。莎士比亚的戏剧是莎士比亚的，而不是其他人的；相比之下，科学发现从根本上说却不属于某一个人。（Humphrey, 1987, Chapter 48）

如果莎士比亚不存在，那么将没人能写出《哈姆雷特》《罗密欧与朱丽叶》和《李尔王》。如果梵高不存在，也就没有人能画出《星月夜》。这样说虽然有些夸张，但也确实说出了点什么。首先，在科学中，大艺术家式的个人贡献是罕见的。不仅如此，从个人贡献的角度来看待科学发展，这本身就想偏了。众所周知，在科学界以及在对诺贝尔奖的竞争中，科学家们为了最先发现某一成果展开激烈的竞赛，这是因为你努力研究出的成果恰恰是别人也可能做出来的，而假如你是某个成果的第二发现者，个人的研究风格可是不会给你加分的。艺术界不存在这种竞争，在那里我们追求的目标是不同的。

有些科学家渴望拥有大量读者，为了取悦读者他们尽量把书写得很有文学性。我第一个想到的就是达尔文的著作。但即便如此，对达尔文而言追求正确性、用所发现的真理来说服读者仍然是第一位的，只要我们把达尔文的《贝格尔号环球航行记》（*Voyage of the Beagle*）和美国著名小说家梅尔维尔的《白鲸》比较一下就能看出来。你当然可以从《白鲸》中学到很多有关鲸和捕鲸的事情，但你也知道梅尔维尔并不是想把它写成一本兼具艺术性和说服力的捕鲸活动纲要。

如果我们记住科学与艺术有着不同的目标，那么我们也可以问科学家一个问题，它与我和哲学同行打趣时问的问题很像：如果梅菲斯特给你以下两个选项，你会选哪一个？

1．通过最先证实了一个发现，你赢得了竞赛，随后获得了诺贝尔奖，人类的科学知识基于这一发现取得了长足的进展，但是就像汉弗莱所说，它并不属于某一个人。（这让我想到了沃森和克里克。如果他俩没有赢得竞赛，莱纳斯·鲍林或者其他人无疑也能很快完成DNA结构方面的有关发现。）

2．你提出了一个原创的理论，它是前人完全无法想象的，甚至你的大名也因此成了我们语言中的一个专有名词。在你去世后的很多年甚至很多世纪中，你的思想激发出了许多有价值的讨论。（这让我想到了笛卡尔的身心二元论、科学家拉马克的进化论、心理学家斯金纳的行为主义理论，还有弗洛伊德的观点，他用幼儿期的性行为与神经官能症来解释一切艺术、音乐和文学。）

其实，笛卡尔的那些野心勃勃的物理学著作可能是选项（2）的一个更好的例子，只不过现在它不是特别有名。但在当时，它的影响力是

相当之大、它的错误又是相当“才华横溢”。正是在笛卡尔的《哲学原理》（*Principia philosophiae*）的刺激下，牛顿写出了改变世界的《自然哲学的数学原理》，牛顿就是为了回应笛卡尔才故意把他的书名起成这样，想明确地表示出他要取而代之。

还有美国语言学家乔姆斯基的语言学理论，从原创标准看，它无疑是达标的。就像人们用“美洲”来冠名“美洲杯”一样，当乔姆斯基为我们打开那片新天地时，完全不存在与他竞争的对手。在以后多年中，最初的理论成果，也就是乔姆斯基在《句法结构》（*Syntactic Structures*）一书中提出的“转化生成”理论已经被广泛弃用，它已被它的各种理论后代所取代，这些理论之间的区别甚大，就像鸵鸟、蜂鸟和信天翁迥异于它们共有的祖先，尽管它们都是从恐龙进化过来的。乔姆斯基在1957年犯了一个有益的错误吗？或者我们不如说他发现了一个伟大的真理？“是的”一词很好的回答了这个问题。

我们崇敬那些犯了有益错误的科学家，想想物理学家沃尔夫冈·泡利是怎么鄙视那些连“犯错都算不上”的理论家们的。如果强迫你选的话，你是想当正确的第一发现者还是想当充满争议的奠基人呢？这个决定有点难做了，是不是？

75

简单的自我人类学

哲学家应该把反直觉的一切都去掉

帕特里克·哈耶斯（Patrick Hayes）是一位人工智能专家，他曾计划将朴素的流体物理学公理化。他当时想要把这些公理加入到一个机器人的核心信念当中，以便它与人类进行互动，因为人们在日常生活中所依赖的正是朴素物理学，或者说常识物理学。这个想法比他预想的更难实现，为此他还写过一篇有意思的文章《朴素物理学宣言》（*The Naïve Physics Manifesto*，Hayes，1978）。在朴素的流体物理学中，所有违反大众直觉的论断都是要被排除出去的：比如，虹吸现象就是“不可能”的，同理，移液器也是“不存在”的；但是用松软的毛巾把液体吸走是可能的，用水泵从井里抽水也是。记住了一大堆这种“知识”的机器人和我们大多数人一样，在第一次看到虹吸实验时也会大吃一惊。

我倒愿意把哈耶斯的想法叫作复杂的朴素物理学，因为他并没有对这种物理学抱有幻想；他知道，虽然他所要公理化的那套理论是我们日常生活中常用的，但它仍然是错的。这么看来，我们也可以把一项活动叫作公理化的人类学：将民众常说的、认同的东西当作你的公理和定理，并且试着消除它们之间存在的矛盾，把这套公理系统融贯起来。当然啦，你不用费功夫去搞什么调查问卷，因为我们既然能像任何正常人一样知晓朴素流体物理学，那么咱们也只需问问自己就能完成这个人类学问卷了，所以它也叫作公理化的自我人类学。(102)

下面可以比较一下哈耶斯的设想与“分析的形而上学”中的哲学设想，后者常常给我一种天真的朴素人类学的印象，因为做这种研究的人似乎深信他们的工作是在获得某种真的东西，而不仅仅是在人类的某特殊群体中被相信为真的东西。然而，他们的工作却似乎是相同的：你收集好你们共享的直觉，让它们在共通的直觉泵中得到检验和激发；再基于“承认下来的”，有价值的、理想的、公理般的原则，努力把最终的数据揉进一个融贯的“理论”。我曾问过许多分析的形而上学家，我问他们

能否把其事业与他们部落的天真朴素人类学区分出来，我至今还没有收到一个令人信服的回答。

与之相反，还有一种复杂的朴素人类学（既包括自我的又不是自我的），它对自身其中的定理是否可信留有判断，而这往往是可行的，也是有价值的。我建议那些用分析的方法做形而上学的人应该把他们的工作重心转移到这项事业中来，因为他们无需大幅调整自己原来的工作方法，只需要重新认识自己存在的理由：他们必须收起自己那副自命不凡的样子，意识到他们的研究只不过是对常识映像进行的一次初步侦查，像人类学家研究异域文化时所做的那样，把对这些信念的信与不信悬置起来：让我们暂时假装认为这些原住民是正确的，看看接下来会发生什么。以我对哲学的理解来看，哲学中至少有一大部分工作都是在反复协调常识映像与科学映像，对于哲学家们来说，在把日常的观点进行理论化并嵌入他们的理论大厦之前，最好先把这些观点好好分析一下。

复杂的朴素人类学的一大特点就是，它对违反直觉的发现保持开放的态度。如果你做的只是朴素人类学，那么你就会把（对原住民来说）反直觉的东西当作你的对手；而当你改变方式，开始思考朴素“理论”在哪些方面是正确的，那么反直觉的论断就不再是你的障碍了，有时它反而是你取得重大进展的标志。毕竟，在科学中，违反直觉的发现是何其珍贵啊。

自我人类学的一个缺陷在于，一个人的直觉容易被他的理论偏好所扭曲。语言学家早就发现了这一点：由于他们太专注于自己的理论，他们自己的语言直觉就不再可靠了。我们能说“小孩男人女人，唱歌吃饭跑步”吗？我的“耳朵”是不是已经被我的从句理论欺骗了？他们原初的直觉已经被太多的理论所污染，他们承认自己必须找那些不是语言学家的人来问一问。最近，哲学家们开始重视这一点，掀起一股所谓实验哲学的新风（Knoke & Nichols, 2008）。这项运动还处在起步阶段，虽然这些先驱者们的努力现在还不引人关注，但这至少表明哲学家们已经开始接受这样一个事实：他们不能再因为一些命题对于他们自己来说是显而易见的，就断言它们显然是真的。同样，如果哈耶斯费心进行一番抽样调查，而不是将他自己的想法当作大众看法的模板，或许他就会对调查出的常识物理学的主要原则感到惊讶呢。

因此，这种复杂的朴素人类学是一项伟大的工程。哲学家们应该严肃地对待这项事业，并把它视作对常识或常识映像进行的一番调查，只有经过这番调查，我们才能开始建构关于知识、正义、真、善、美、时

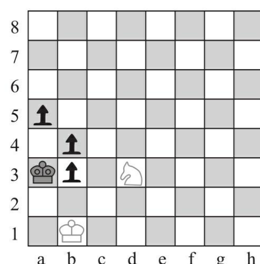
间、因果性等等诸如此类的理论，也只有这样才能保证我们的分析与论证找对了靶子，切中了民众与科学界真正关心的东西。这样一番系统的调查将会给我们提供一个目录式的东西，给未经修正的概念领域编一个目录。这些概念给理论家们带来了一系列困难，如果你乐意的话，可以把它们叫作常识映像的形而上学。当哲学家们试图不断协调那些最新的科学观念时请务必从这里开始，而用这样一幅更严谨的观念地图来取代肉眼目测是有益无害的。

或许有人会说，我们还需要改革，让科学哲学从摇椅上的幻想变为与现实科学的严肃合作，让科学哲学家们真的决心从当前的科学研究内部得到有关科学的知识。如果我们从这个角度思考哲学的任务，就会发现我们的哲学期刊中充斥着大量没有说到点上的苦思冥想与反复纠结，要么就是兜售反例、打破直觉，它们最多算是一次为那些“显而易见”的观念寻找共识的尝试罢了。

76

“像棋”中的二阶真理
若不值得做，亦不值得做好

残局如下，白棋如何在两步之内将死对方？(103)



答案：第一步，白马走到b2，黑兵走到a4（黑棋只能这么走）。第二步，白马走到c4，将军。

《波士顿环球报》刊载了这样一个残局，让我大吃了一惊，我一直以为在国际象棋中无法“单马定孤将”是已经被证明了的呀，可是我错了。就像戴维·米希亚洛维思奇（David Misialowski）最近给我写的一封电子邮件中说的，人们确实已经证明了当你只有一马一王且对方只有一王时无法将死对方的王。所以，“‘当你只有一马一王时无法将死对方的王’这一命题不是国际象棋中的真理”这一事实是国际象棋中的高阶真理。

哲学一般被认为是一门先验学科，就像数学一样，或者说它至少是以某种先验的方法论为核心。我们可以从两个角度来看待这一点。一方面，哲学家们能以此为借口逃避枯燥的实验室工作，不用去做田野调查，也不用学着收集数据、学习统计学，不用学地理、历史、外语、经验科学等等，因为只有这样他们才有大把的时间来磨练自己的哲学技巧。而另一方面，就像人们常说的：万事万物中都包含着哲理。这可未必是一件幸事。在这一节中，我要对你们这些想以哲学为业的年轻读者们（但愿你们当中有人这么想）提出一点忠告：自由与抽象性恰恰可能是哲学的一个弱点。而对于哲学的门外汉们而言，这一节也能帮助你们了解哲学中的某些习惯和误区。

我们可以看一看先验真理中的范本，即国际象棋真理。人们下国际象棋就是一个经验事实，我们关于下国际象棋的经验事实有很多，它包含了几个世纪以来人们下棋的方式，包含了精雕细刻的棋子在镶饰的棋盘上的所有走法，等等。但这些经验事实的知识在我们研究国际象棋的先验真理时并不能发挥作用，而这样的先验真理在国际象棋中其实大量存在。我们只需要知道国际象棋游戏的规则就能去探索先验真理。国际象棋开局就足足有20种走法（走兵的16种，走马的4种），一王一相无法将死孤王，一王一马也不行，如此等等。

但先验国际象棋真理的推导并不都是那么容易。在国际象棋规则内，想要证明什么可能而什么不可能是一项非常复杂的工作，出错更是在所难免的。比如，几年前，一个计算机国际象棋程序发现了一种将杀网（相当于一串必胜的走法），在它执行过程中，竟有200余步是不吃子的。这一发现打破了国际象棋中的一个古老“定理”，并且迫使人们改变了国际象棋规则。曾经人们认定，在50步之内如果双方都没有吃子，那么这盘棋就算打平了（僵局），但是自从有了这个必胜的、牢不可破的将杀网的存在，继续坚持“50步内未吃子即是平局”的规则就变得没有道理了。在计算机开始下棋之前，没有人能想到“会有”步数如此之多的将杀网存在。思考这些问题是非常有趣的，很多高智商的人都曾致力于研究国际象棋的先验真理体系。

很多哲学项目或问题的研究，就像在推算国际象棋的先验真理。这两种活动都是先以一组大家一致同意而几乎从不讨论的规则为前提，然后再对这些前提所包含的东西进行推算、组织、争论和重新定义。话说到此还看不出什么问题。国际象棋是人类创造的一项深奥而有意义的活动，关于它的价值人们已经有过太多的论述。可是某些哲学研究项目却更像是在推算“像棋”（chmess）的真理[(104)]。“像棋”很像国际象棋，唯一的区别是“像棋”中的王可以一次走两格而不是一格。我刚刚把它发明出来，当然，大家可以深入研究它的可玩性，也许它一点也不好玩，也许它该起个什么别的名字。算了，我不费心研究这些问题了，因为虽然这些问题确实可回答，但并不值得我花时间和精力去探究它们。“像棋”中的先验真理和国际象棋中的一样多，都是无限的，而且一样难以发现。这意味着，如果有人真的投身于“像棋”真理的研究，他也可能会出错，也可能需要被纠正，在这个过程中就开辟出了一块先验研究的新领域——“像棋”中的高阶真理，例如下面这些：

1．琼斯对“像棋”真理*p*的证明（1989）有问题：他并未考

虑如下的可能……

2．史密斯断言（2002）：琼斯证明（1989）错误地预设了布朗引理（1975）为真，而布朗引理最近遭受了加芬克尔的挑战（2002）……

这些不只是儿戏了对吧？事实上，可能会有人向你证明，人们对“像棋”高阶真理的研究成果辉煌呢。我想，心理学家唐纳德·赫布（Donald Hebb）的格言该派上用场了：

若不值得做，亦不值得做好。

如果真要将赫布的格言贯彻到底，可能哲学家们都会觉得他们所参与的某个持久的哲学争论该被叫停了，当然啦，我们是不会同意只关闭一些“手工作坊”的。因为在有些学派看来，我们这个庞大的学科中根本就不存在什么研究，只有一些无用功和费尽心力地互相洗白。民主或许产生不出什么值得注意的东西，可是独裁更可怕，所以还是让我们百花齐放吧。不过要记住的是：假如你想看百花齐放，就要为99朵花终将会枯萎做好准备。其实我想提醒你们的只是：不要把你们自己成长的宝贵岁月浪费在一个保质期很短的研究主题中。哲学潮流来得快去得也快，人们常说“炒得越火的话题凉得越快”，这可能有点道理。

要想检验一项哲学研究是否只是在探索“像棋”的高阶真理，一个好办法是看看除了哲学家之外这个游戏是否真的有人在玩。除了学院中的哲学家之外，还有人“关心”琼斯的反例能否驳倒史密斯的证明吗？把这些东西讲给入行未深的本科生们听，是一种不错的检验方法。如果他们“听不懂”，你就该好好想想是不是自己在被那个自娱自乐的专家团体牵着鼻子走，最后掉进了人造的陷阱里去了。

下面我们就看看人们是怎么掉进陷阱的。哲学工作不是一项自然的活动，越聪明的人越有可能怀疑自己是否有研究哲学的天赋、是否理解哲学、是否在“以正确的方式”研究哲学，甚至最初就会怀疑以它为业是否值得。这不，聪明的琼斯就是一位对进入哲学系感到不安的本科生。他对布朗教授开设的讨论班深深着迷，于是试着写了一篇有关热门话题H的论文，布朗教授给了他一个A。“你真的很有天赋啊，琼斯，”布朗教授说道，这让琼斯看到了一项似乎值得奋斗终生的事业。于是琼斯开始在与其他同样志向远大的年轻人的互动中，努力地学习这个游戏的规则。他们相互打气道：“嘿，我们可真擅长哲学啊！”“为了论证的方

便”，他们逐渐不再怀疑那些使这个游戏得以开展所预设的前提，开始出书了。

所以，不要指望你的研究生同学或是你喜欢的教授能告诉你，你是否只是在探索“像棋”的高阶真理。因为他们都是这项事业既得利益的维护者。这种问题在其他领域中我们也会碰到，甚至可能还更难破除。有时一位技术纯熟的实验员也会陷入窘境，他成天只能在开销不菲的实验室中往表单里填入一些不会再有谁关心的实验数据。他们又该怎么办呢？难道要把那些昂贵的设备扔掉不成？这是个棘手的问题。而对于哲学家来说，重装上阵其实要简单很多，也便宜很多。毕竟，一般来说我们所接受的“训练”算不得什么高科技。我们主要是在各种文献中进行学习，学习那些久经考验的招数。因此我们需要躲避的陷阱仅仅是这样的：当你看到了一位著名学者在他写的东西里提出了一个站不住脚或者值得怀疑的断言，比如，你发现了一个容易攻击的目标，教授一篇不失巧妙的论文中有一个不严谨之处，你打算把这个目标靶子当成是在学术界闪亮登台的一个不错处女作。于是你大胆尝试，参与到与他人的辩论中。现在你可要小心了，因为当你们相互引用，再对回应进行回应时，你就已经成了一个崭露头角的专家了，“精通”于处理有关“如何处理教授这个小纰漏的回应”这方面的问题。还要记住的是，如果教授没有大胆提出那一论点的话，他根本不可能一下吸引那么多注意；更何况，脱颖而出的诱惑力，并不只限于想在哲学界搏出位的年轻人。

有些人安于找到一个志趣相投的朋友圈子，在这些聪明人中间享受发现的乐趣、合作的喜悦，还有和他人达成一致时的满足感。就像哲学家约翰·奥斯汀（John　Austin）在他已出版的一篇演讲中说的那样：不用担心大家一起做的工作是不是有价值的。如果从事这项工作的人足够多，似乎这一现象本身就会变成这项工作值得研究的证据。哲学家波顿·德雷本（Burton　Dreben）曾在哈佛大学对研究生们这样说道：“哲学是垃圾，但垃圾的历史则是学问。”有些垃圾比别的垃圾更重要，但我们仍然很难决定该把哪些垃圾当成学问。奥斯汀演讲录的另外一篇演讲中有一句话，堪称暗讽的杰作：

> 来听我讲座的听众一般都不是来听重要的事的，如果今天来的诸位中有人想听重要之事，事先声明：我欠你们一个结论。（Austin, 1961, p.179）

奥斯汀是一位才华横溢的哲学家，有许多有前途的哲学家都曾围绕

在他身边，当听到奥斯汀说这番话时他们肯定也会在台下会心一笑吧。可如今，他们当中的大部分人已经消失得无影无踪了，他们关于日常语言哲学的著作（“哇，写得真聪明”）也在出版后不久就被人们理所当然地遗忘了。这样的事情屡见不鲜。

那么我们应该怎么办呢？之前我提到过一种检验方法，即，看看不做哲学的人或者聪明的本科生是否会喜欢你的研究，但它只算是一个警告标志，并不绝对。某些令人望而生畏的深奥的哲学问题也是值得研究的，即便外行们对它不为所动。我当然也不会阻止你在研究时无视别人是否觉得它有意义、是否重要。相反，哲学中那些伟大的创见最初总会遭到怀疑和无情的嘲弄，你的脚步不应被它们阻挡。我想说的只是，请不要只因为自己那些聪明的同行觉得你的工作不容忽视——就像你也认为他们的工作很有意义一样——就沾沾自喜地以某个思想流派而自居。或许，你们只是在相互欺骗罢了。

77

只需关注那10%的精品
哲学领域的史特金定律

如果哲学像其他领域一样，也适用史特金定律，那么哲学领域中的哪些作品会被我算作精品呢？首先，“经典”绝对是经典的一个理由。从柏拉图到罗素，这些哲学史教程的标配甚至已经通过了几个世纪的检验，而且研究他们的一些优秀二手文献也很有价值。你即便不了解那些背景知识，直接去读亚里士多德、康德、尼采的著作也会获得一些或是很多启发，但若是能接受一些专家的指导，你将从他们身上学到更多，因为这些专家早已把毕生的心血都投入到对那些伟大思想家的研究中去了。

并不是所有哲学史家都抱着相同的目的和态度写作，我就不认为我们有任何理由剥夺别人写作哲学史的资格。有些人坚持把那些思想家放在他们写作时的历史背景中去理解，也就是说，他们认为如果你真的想要理解笛卡尔就该深入了解17世纪的科学，如果你真的想理解洛克和休谟就应该阅读17或18世纪政治史，当然，有很多比他们稍逊色的同代学者你也应该多加关注。

但我们为什么要在这些失败者身上花心思呢？我可以举一个好例子来说说为什么。当我去欧洲的博物馆参观，看见堆积如山的二流作品之后，我才发现自己从未真正欣赏到16～17世纪的那些名画之美。如果你所看的全部只是精品，比如你在导论课上、在顶级博物馆里看到的那些东西，那么你反而很难发现它们的精妙之处。你知道好图书馆和一座伟大的图书馆之间的区别吗？好图书馆里收藏得全是好书，而伟大的图书馆里收藏着所有的书。如果你真想理解一位伟大的思想家，你必须花时间读读与他同代的或前辈学者的作品，这些作品常常被大师们的作品掩盖。

另外一些专家只是稍稍触及历史背景，而主要关心如何将这些伟大作品中的观点与今天的问题联系起来。莱布尼兹毕竟不是为了写一篇17世纪理性主义的代表作而写《单子论》的，他是为了揭示真理而写。说

到底，只有当你追问这些哲学家所讲的是否正确时，才算得上严肃地对待他们。哲学生和哲学教授们有时会忘记这一点，他们关注于怎么给那些哲学家分类，忙于“比较和对比”哲学家之间的关系，就像是在回答考题。有时整个哲学系都会陷入这类愿景中无法自拔。可是这不是哲学，这是哲学鉴赏。我是这样帮我的学生们打破这一陋习的：

> 一个人无意中发现了一个惊天大秘密：比如有人图谋炸毁自由女神像，或者破坏国家电网。于是他疯狂地工作，收集和整理证据，尽力发挥文采写了一封长信。他把信寄给了警察、联邦调查局、《纽约时报》和美国有线电视新闻网。随后得到的回复是这样的：“哈，又是一篇巧妙的后9·11阴谋论”“引人入胜、内容相当可信且能自圆其说、文笔甚佳”“它让我想起了小说家唐·德里罗（Don De Lillo），当中亦有托马斯·品钦（Thomas Pynchon）写作风格的影子。”啊啊啊啊啊！注意！我正在告诉你们实情！请尊重你正在解读的哲学家，所以，当你读他们写的每一句、每一段时，请你多问问自己：“我相信这些话吗？如果不相信又是因为什么？”

除了纯哲学之外，还有科学哲学、数学哲学、逻辑学哲学、生物学哲学、心理学哲学、经济哲学、政治哲学，在这些领域中也有伟大的著作。虽然几乎不存在所谓化学哲学、天文学哲学、地理哲学或者工程学哲学，但是一些出色的哲学工作也会处理这些领域中存在的概念问题。还有就是伦理学了。1971年，约翰·罗尔斯（John Rawls）出版了《正义论》（*A Theory of Justice*），这部杰作开创了一个硕果累累的新时代，哲学家们开始从新的角度处理传统的伦理问题，他们开始关注社会科学，尤其是经济学、社会学，甚至还有生物学和心理学。多亏了罗尔斯，伦理学家们的游戏被抬高到了一个新的水平，产生了许多有价值的哲学研究，这些工作值得、也的确得到了其他学科的研究者们的关注，甚至包括政治家和社会批评家。

最后，还有一些哲学家完全不跨学科，也只稍稍借助一点哲学史，他们专攻某些哲学家工作中出现的当代的问题，而这些哲学家也专攻其他的当代哲学家工作中出现的当代的问题。其中一些我在上一节里提过了，他们没能遵从“赫布定律”：若不值得做，亦不值得做好。但其中还有一些工作是非常优秀的，并且很有价值。在这本书里我提到过许多当代哲学家，如果我觉得他们的想法不值得严肃对待的话，我也就不会提

到他们了，尤其是当我说他们想错了的时候。除了我的那些靶子之外，我还欣赏很多哲学家的工作，但我不想傻傻地去列个名单。

在我的学术生涯中，我曾多次轻信同行的判断，他们告诉我别读某人的书了，写得又蠢又垃圾。过了一段时间我才发现自己被误导了，竟忽视了一位有创造性的思想家，同行的偏见害得我没能及时吸收这些人的思想。而且，我痛苦地意识到，即使是一些没能成功进入我阅读列表的哲学家，当我读到他们时也会发现他们的有趣之处。所以，请把这本书当作一本开放式的导论，它介绍了某些哲学工作的方式，如果你认为它们有意义，就可以把它们作为跳板，以自己的方式探索那些被历代思想家们反复研究的问题吧。

结语　加倍努力地使用这些工具吧

“太不可思议了！”，当某些人发现意识的“秘密”时如是说；“太不可思议了！”，当某人发现这个星球在30亿年前不借助任何造物主之手就产生了生命时如是说。每当我听到这种话，总会忍不住想说：“好吧，对你来说这当然不可思议了。因为你自己把思考工具扔在一边，几乎从来不碰。”这让我想起了英国生物学家威廉·贝特森曾坚定地声称基因的物质基础是不可想象的。而今天，即使对学校里的孩子来说，DNA也是很容易理解的。这并不是因为他们比贝特森高明多少，而是因为经过一个世纪的设计和改良，我们的思考工具已经足以快刀斩乱麻，做出准确的判断。当然，还有一些人确实不愿构想这些事情，哪怕连一个解释都不肯给出，他们想要保护神秘之物，担心解释会令宝藏消失。

他们发现，当别人想要刨根问底时，“上帝以神秘的方式起作用”是一种方便的反思考工具。它暗示出发问者的傲慢和不自量力，把好奇心瞬间消除。以前这个工具很好用，现在在一些社群中它也依然好用，因为对于那些人来说，忽视科学只不过是一个微不足道的缺点。我想，我们实在不该把这种“虔诚”的想法当作智慧了，我们应该意识到它俨然是一种预防性的宣教。我们应该积极地回应道：“噢，太好了！我爱神秘的东西。让我们瞧瞧能不能把它也解决掉吧。你们有什么点子吗？”

构想新事物是一件困难的事，这不只是“在头脑中构建出一个观点，迅速浏览一遍然后表示赞成”而已。如果我们研究得更加努力，对我们来说不可构想的东西或许现在已经显然可以构想了。有时我们可以自信地宣称某些东西确实不可构想，比如最大的质数、内角和大于180度的平面三角形、已婚的单身汉等等。与其说是因为我们发现我们无法构想它们，不如说是因为我们发现自己已经如此清楚、如此彻底地设想出它们以何种方式构成，所以它们的不可构想性本身就被我们非常清楚地构想出来了。

我们仍未完全成功地构想出意义如何在物质世界中存在、生命如何诞生与进化、意识如何工作以及我们是否拥有自由意志，但我们已经取得了进展：与过去相比，现如今我们正在用更恰当的方式提出问题、解决问题。我们正在寻找答案。

附录一　未写入本书正文中的几个直觉泵

读过本书草稿的朋友们曾对我表达过他们的惊讶和失望，因为几个我最著名的直觉泵并没有出现在本书的定稿中。实际上，最后被删掉的有几十个呢，其中有些是我的最爱。这种情况并不常见，我想应该解释一下我这么做的必要性。

“我在哪儿”可能是我最著名的一个直觉泵了，但也正因如此才应该把它删掉。1978，它最早出现于《头脑风暴》一书，后来在《心我论》中也谈到了它。这些书已经被翻译为了十几种语言，而这个直觉泵也常常被编入各种选集。电影《大脑的受害者》（*Victim of the Brain*）中有30分钟的情节也是根据它改编的，我还扮演了第二个丹尼特的身体。1981年BBC拍摄的关于意识与大脑的纪录片中它亦有出镜。著名木偶师林恩·杰弗里斯（Lynn Jeffries）还把它改编成了爪哇皮影戏，1984年在哈佛大学勒布剧院里演出过。你肯定能很容易就能在网上查到它和关于它的大量评论。

“沙基披萨店的民谣”（Dennett, 1982a）消解了“从物模态与从言模态的信念”所包含的一套预设。对于某些哲学家的意向性思想而言，这套预设曾占据统治地位，而对于其他人来说则不是这样。所以，如果要把它写进这本书里，我就得先诱导诸位产生一些错误的直觉，然后才能用这个直觉泵把你们治好。有些哲学家必须了解这个直觉泵，而对于其他人嘛，继续幸福地无视它也不会有丝毫的损失。

在进化论方面我也忍痛割爱，删去了我最喜欢的一个思考工具，即哲学家皮特·戈弗雷-史密斯发明的“达尔文的空间”，它是我见过的把多维空间作为思考工具引入哲学的一次最佳尝试。但遗憾的是，我们必须仔细回顾进化论并介绍大量的生物学事实才能让这个直觉泵有效地工作起来。我用更技术化的术语给其他的哲学家讲解过，并且在我为戈弗雷-史密斯的著作《群体思维与自然选择》（*Population Thinking and Natural Selection*）所写的那篇评论——《小人儿法则》（*Homunculi Rule*）中也展示过它有多么好用。如果大家感兴趣，还可以参见戈弗雷-史密斯对我的回应（Godfrey-Smith, 2010）。

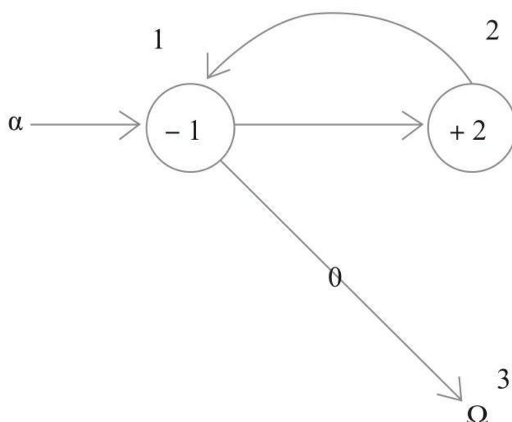
在我的论文《蒯因感受质》（*Quining Qualia*, Dennett, 1988a）里

包含了多达14个直觉泵，它们是我为了澄清并消除那个混乱得无可救药的哲学概念“感受质”而发明出来的。为了介绍感受质这个概念以及它所带来的问题，本书收录的“花椰菜的诅咒”即是其中一个直觉泵。对于那些仍然把感受质概念看成一个好概念的人，或许该读读我的这篇论文。很多文集都收录了这篇论文，网上也能找到，各种语言的都有。在我写的《做个好梦》（*Sweet Dreams*, Dennett, 2005b）一书中包含了与“感受质”有关的另外一些论证和直觉泵。关于意识的直觉泵还有《为什么你无法造一台可以感受疼痛的计算机》（*Why You Can't Make a Computer That Feels Pain*, Dennett, 1978c）中的“箭毒致忘剂”（curare-cum-amnestic）、《机器玛丽知道什么》（*What RoboMary Knows*, Dennett, 2007d）中的“沼泽玛丽与机器玛丽”（Swamp Mary and RoboMary）、《意识的解释》中的意识模型“奥威尔式与斯大林式”等等。我需要花很多篇幅才能把它们讲清楚，可是我不希望这本书写得过长。

被删去的还有我在《打破魔咒》（*Breaking the Spell*, Dennett, 2006a）中介绍过的很多有关宗教的直觉泵，以及《科学与宗教：它们相容吗？》（*Science and Religion: Are They Compatible*, Dennett and Plantinga, 2011）中“超人动了波基斯页岩”的例子。

附录二　寄存器机练习题的答案

练习1



程序1

a．运用程序1，想一下寄存器机需要几步可以计算出2＋5＝7　　？（“结束”也算是一步。）

答：6个步骤：3个减量，2个增量，一步结束（最后一步是遇0跳转）。

b．要计算出5＋2＝7需要几步完成？

答：12个步骤：6个减量，5个增量，一步结束。

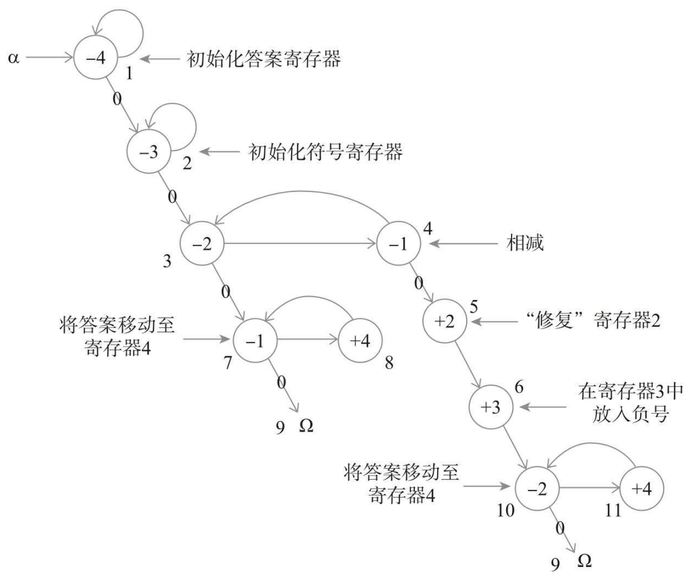
（你从中可以得出什么结论？）

答：内容按不同顺序相加会导致很大的差异，所以你可能会认为，最好设定一条规则，总是将较小的数字放入寄存器1；但如果你要首先从两个数字中区别出那个较小的，这会比你完成相加花费更多步骤！

练习2

a．请为这个流程图写出它的RAP程序。（请注意，由于程序有不同的分支，你可以采用不同的顺序为各步骤编号，只要“下一步”对应的步

骤编号填写正确即可。）



步骤	指令	寄存器	下一步	跳转至步骤
1	减量	4	1	2
2	减量	3	2	3
3	减量	2	4	7
4	减量	1	3	5
5	增量	2	6	
6	增量	3	10	
7	减量	1	8	9
8	增量	4	7	
9	结束			
10	减量	2	11	9
11	增量	4	10	

b．如果从3中减3，或者从4中减4，这个程序会出现哪些不同？

答：程序在寄存器4那里遇0中止。

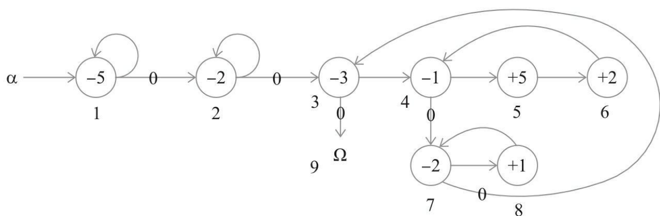
c．将寄存器3在第三步之前就清零而不是推延到完成第四步之后，这样会避免哪些可能的错误？

答：如果寄存器1和寄存器2里最初放置的都是0，那么在程序结束

后，答案很可能无意义（符号寄存器里出现-0或除0和1之外的其他数字）。

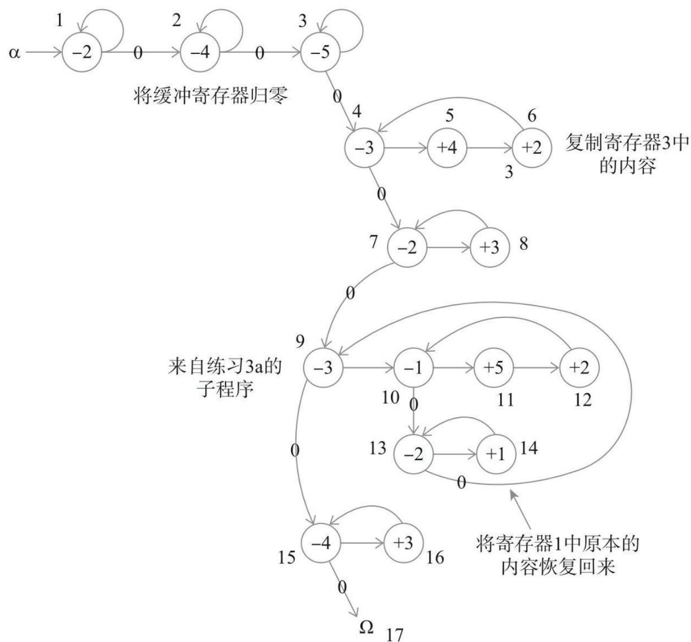
练习3

a．画一个流程图或者写一个RAP程序，让寄存器1中的数字与寄存器3中的数字相乘，将答案显示在寄存器5中。



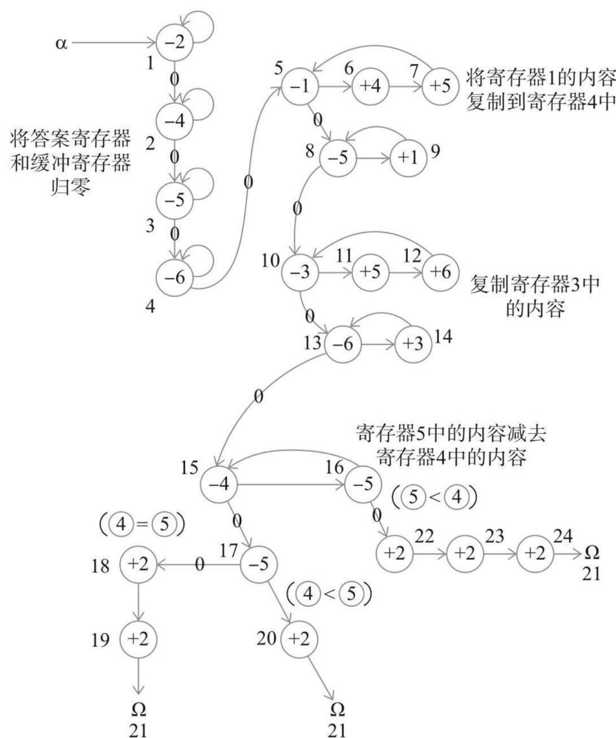
步骤	指令	寄存器	下一步	跳转至步骤
1	减量	5	1	2
2	减量	2	2	3
3	减量	3	4	9
4	减量	1	5	7
5	增量	5	6	
6	增量	2	4	
7	减量	2	8	3
8	增量	1	7	
9	结束			

b．（可选）利用复制和移动改进你上面写出的乘法编程，使寄存器1与寄存器3在程序停止后都能保持原有的内容，方便你在运算完成后清楚地检查输出和输入的内容正确与否。



步骤	指令	寄存器	下一步	跳转至步骤
1	减量	2	1	2
2	减量	4	2	3
3	减量	5	3	4
4	减量	3	5	7
5	增量	4	6	
6	增量	2	4	
7	减量	2	8	9
8	增量	3	7	
9	减量	3	10	15
10	减量	1	11	13
11	增量	5	12	
12	增量	2	10	
13	减量	2	14	9
14	增量	1	13	
15	减量	4	16	17
16	增量	3	15	
17	结束			

c．（可选）画一个流程图或者写一个RAP程序，要求它能识别寄存器1与寄存器3中的内容（同时保证这些内容不受损坏），将内容较大的寄存器地址（1或3）显示于寄存器2；当两个寄存器中的内容相等时，寄存器2中显示2。（程序完成后，寄存器1和寄存器3中的内容要保持不变，寄存器2显示出所含数字较大的寄存器地址，当内容大小相等时，寄存器2将报出数字2。）



步骤	指令	寄存器	下一步	跳转至步骤
1	减量	2	1	2
2	减量	4	2	3
3	减量	5	3	4
4	减量	6	4	5
5	减量	1	6	8
6	增量	4	7	
7	增量	5	5	
8	减量	5	9	10
9	增量	1	8	
10	减量	3	11	13
11	增量	5	12	
12	增量	6	10	
13	减量	6	14	15
14	增量	3	13	
15	减量	4	16	17
16	减量	5	15	22
17	减量	5	20	18
18	增量	2	19	
19	增量	2	21	
20	增量	2	21	
21	结束			
22	增量	2	23	
23	增量	2	24	
24	增量	2	21	

练习4（可选）

写出一个流程图及它的RAP程序，用来将寄存器机转变成一台便携式计算器。

a．寄存器2用来执行运算：

0＝相加

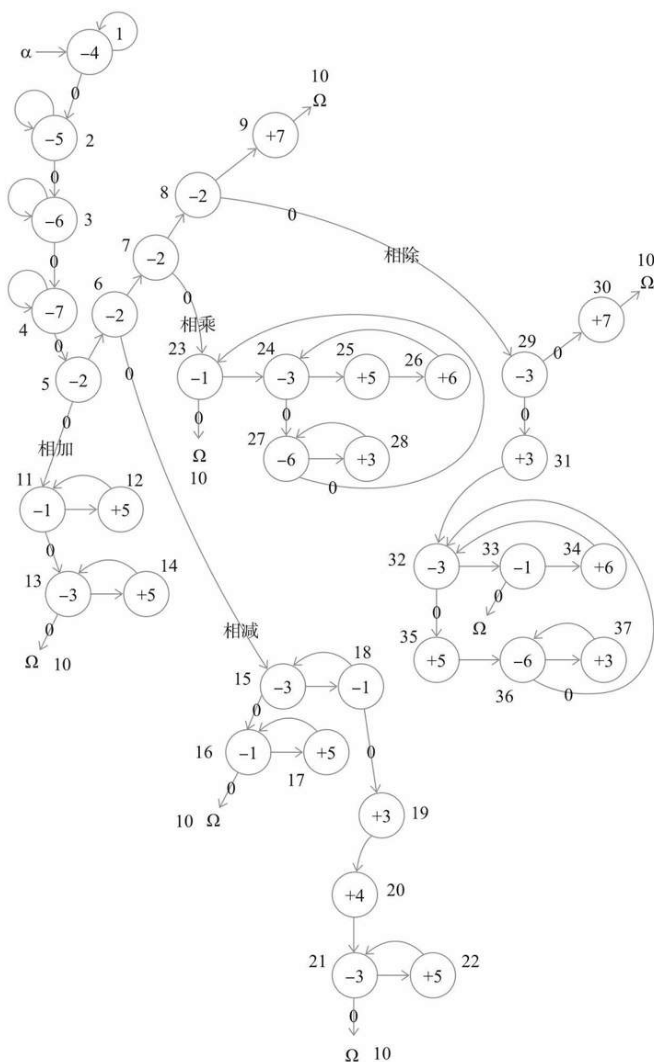
1＝相减

2＝相乘

3＝相除

b．在寄存器1和3中放入参与运算的数据。

那么“306”就表示“3 + 6”，“513”表示“5-3”，“425”表示“4×5”，“933”表示“9÷3”。然后，我们用寄存器4、5、6、7来显示操作结果：寄存器4用于表示符号，其中0代表＋，1代表-；寄存器5显示答案的数值；寄存器6用于存放除法结果的余数；寄存器7表示警示，即对输入的内容进行报错，例如除数为0，或者寄存器2中出现了没有定义的操作的情况。



步骤	指令	寄存器	下一步	跳转至步骤
1	减量	4	1	2
2	减量	5	2	3
3	减量	6	3	4
4	减量	7	4	5
5	减量	2	6	11
6	减量	2	7	15
7	减量	2	8	23
8	减量	2	9	28
9	增量	7	10	
10	结束			
11	减量	1	12	13
12	增量	5	11	
13	减量	3	14	10
14	增量	5	13	
15	减量	3	18	16
16	减量	1	17	10
17	增量	5	16	
18	减量	1	15	19
19	增量	3	20	
20	增量	4	21	
21	减量	3	22	10
22	增量	5	21	

步骤	指令	寄存器	下一步	跳转至步骤
23	减量	1	24	10
24	减量	3	24	27
25	增量	5	26	
26	增量	6	24	
27	减量	6	28	23
28	增量	3	27	
29	减量	3	31	30
30	增量	7	10	
31	增量	3	32	
32	减量	3	33	35
33	减量	1	34	10
34	增量	6	32	
35	增量	5	36	
36	减量	6	37	32
37	增量	3	36	

素材来源

引言

“古怪的狱卒”和“垃圾箱里的珠宝”取自*Elbow Room*（Dennett, 1984a）中刻画的案例，并做了一些修改。

第一部分　通用思考工具

1．“犯错儿”以“How to Make Mistakes”（Dennett, 1995b）为开头。

2．“归谬法”之前从未发表过。

3．“拉波波特法则”，这一法则是我在针对理查德·道金斯的*The God Delusion*发表的一些思考中首次提出的，发表于*Free Inquiry*（Dennett, 2007a）。

4．“史特金定律”，这一定律在“Holding a Mirror up to Dupré”（Dennett, 2004）中有相关讨论。

5．“奥卡姆剃刀”之前从未发表过。

6．“奥卡姆扫把”只是在很多网站上被归为西德尼·布伦纳发明的新概念。而他本人似乎并没有发表过有关这一概念的文章。

7．“外行做媒”之前从未发表过。

8．“跳出系统”借鉴了“I Could Not Have Done Otherwise—So What?”（Dennett, 1984b）中的素材。

9．“古尔德的3种思考工具”取自“Confusion over Evolution: An Exchange”（Dennett, 1993）和“Shall We Tango? No, But Thanks for Asking”（Dennett, 2011b）。

10．“小心‘当然’这个词”出自“Get Real”（Dennett, 1994a）。

11．“反问”之前从未发表过。

12．“什么是‘深马’”取材于“With a Little Help from My Friends”（Dennett, 2000）。

第二部分　关于意义和内容的思考工具

13．“特拉法尔加广场上的谋杀案”取自“Three Kinds of Intentional Psychology”（Dennett, 1981）。

14．“生活在克利夫兰的一位兄长”取自“Brain Writing and Mind Reading”（Dennett, 1975）。

15．“‘爸爸是名医生’”取自*Content and Consciousness*（Dennett, 1969）。

16．“常识映像和科学映像”一节包含从“Expecting Ourselves to Expect”（Dennett，即将发表）、“Sakes and Dints”（Dennett, 2012）和“Kinds of Things”（Dennett，即将发表）中提取的材料。有关食蚁兽和鸟类特性的讨论取自*Elbow Room*（Dennett, 1984a）。

17．“常识心理”取自“Three Kinds of Intentional Psychology”（Dennett, 1981）。

18．“意向立场”取自“Intentional Systems”（Dennett, 1971）和*The Intentional Stance*（Dennett, 1987），也在其他很多文章和书籍中有所讨论。

19．“人与‘次人’的区别”取自*Content and Consciousness*（Dennett, 1969）。

20．“大脑中的小人儿委员会”在Brainstorms（Dennett, 1978）和Edge.org（Dennett, 2008）的一篇文章中都有提到，后来发表在*What Have You Changed Your Mind About*（Brockman, 2009）。

21．“近似算子”之前从未发表过，但其讨论的主题是从“Turing's ‘Strange Inversion of Reason’”（Dennett，即将发表）中发展而来的。

22．“神奇组织”，这个概念最早出现于我的文章“Natural Freedom”（Dennett, 2005a）。感谢克里斯托夫·科赫，让我注意到了威廉·贝特森的引文。

23．“身陷机器人控制室”取自“Current Issues in the Philosophy of Mind”（Dennett, 1978b）。

第三部分　关于计算机的思考工具

24．“计算机施展魔法的7个秘密”是从塔夫茨大学设立的计算机科学导论课程中发展而来的，这门课程由我、乔治·史密斯和戴维·艾尔斯（David Isles）一同教授，在计算机科学专业本科生史蒂夫·巴尼（Steve Barney）的大力帮助下顺利完成。期间，史蒂夫为我们的课程设计了模拟计算机AESOP，他也因此被斯隆基金会看中，从而成为了其资助下的塔夫茨课程软件工作室的首席程序员。深受AESOP的启发，1986年我用Logo语言完成了第一版RodRego的编程，将它命名为Rego，130页插图中那个尖头小矮人就是Logo语言里海龟绘图的古老痕迹。之后，罗德·达·席尔瓦（Rod Da Silva）将Rego改编成更便捷、更可靠的版本RodRego，并在课程软件工作室的尼克拉·施韦尔特纳（Nikolai Shvertner）那里不断发展完善。

25．“虚拟机”，我在很多地方都讨论过它们，比如在课程软件工作室的成立宣言“Notes on Prosthetic Imagination”（Dennett, 1982d）、“The Practical Requirements for Making a Conscious Robot”（Dennett, 1994b）中，以及在我的著作*Consciousness Explained*（Dennett, 1991a）里。

26．“算法”取材于*Darwin's Dangerous Idea*（Dennett, 1995a）。

27．“让电梯实现自动控制”是我2007年在哥本哈根大学参加研讨会时、在“Varieties of Content”（Dennett, 2007c）中首次提出的。未曾发表。

“第三部分小结”中的部分思考来自“Turing's ‘Strange Inversion of Reasoning’”（Dennett，即将发表）。

第四部分　更多关于意义的思考工具

28．“红发人那事儿”取自“Things about Things”（Dennett, 2001e）。

29．“彷徨的双币机、孪生地球以及巨型机器人”是根据“Evolution, Error and Intentionality”（Dennett, 1988c）修改完成的。

30．“彻底翻译与蒯因式填字游戏”最早出现于“With a Little Help from My Friends”（Dennett, 2000）。

31．“语义引擎和句法引擎”，根据“Three Kinds of Intentional Psychology”（Dennett, 1981）修改完成。

32．“沼泽人遇上母牛鲨”中包含的很多素材都是根据“Features of Intentional Action”（Dennett, 1968）和“Get Real”（Dennett, 1994a）中的内容修改得来的。

33．“两个黑盒子”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

第五部分　关于进化论的思考工具

34．“万能酸”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

35．“孟德尔图书馆”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

36．“什么是基因”这篇新作，深受理查德·道金斯的*The Ancestor's Tale*（2004）这部著作的启发，并引用了书中的大量原文。

37．“生命之树”记录了我的新想法，其中的配图来自伦纳德·艾森伯格（Leonard Eisenberg）的创作作品。

38．“自然选择VS智能设计”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

39．“小布谷鸟为什么会把宿主的蛋推出鸟窝”，根据“Darwin's ‘Strange Inversion of Reasoning’”（Dennett, 2009a）中的相关内容修改完成。

40．“白蚁城堡是谁设计的”在“Intentional Systems in Cognitive Ethology: The 'Panglossian Paradigm' Defended”（Dennett, 1983）中首次出现，在其他很多文章中也都有讨论。

41．“蝉知道素数的含义吗”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

42．“瞪羚的弹跃”是新内容。

43．“世界上第一只哺乳动物是什么”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

44．“物种形成于何时”，根据“The Multiple Drafts Model”（Dennett and Akins, 2008）中的相关内容修改完成。

45．“回溯性加冕”，根据*Darwin's Dangerous Idea*（1995a）中的相关内容修改完成。

46．“循环”是根据“Cycles”（Dennett, 2011a）修改完成的，这是我对于Edge网站2011年的问题“What Scientific Concept Would Improve Everybody's Cognitive Toolkit?”的答案，会在之后的*This Will Make You Smarter*（Brockman, 2012）中出版公布。

47．“青蛙的眼睛究竟将什么告诉了青蛙的大脑”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

48．“穿越巴别图书馆”，根据*Darwin's Dangerous Idea*（Dennett, 1995a）中的相关内容修改完成。

49．“谁是《帕姆雷特》的作者”，根据我在美国东部哲学协会发表的主席报告的相关内容修改完成，当时的报告题目是“In Darwin's Wake, Where Am I?”（Dennett, 2001d）。

50．“虚拟旅馆的噪音”，根据“Collision-Detection, Muselot, and Scribble: Some Reflections on Creativity”（Dennett, 2001a）修改完成。

51．“哈尔是赫布和爱丽丝的孩子吗”，根据“Homunculi Rule: Reflections on *Darwinian Populations and Natural Selection* by Peter Godfrey-Smith, Oxford University Press, 2009”（Dennett, 2010）中的相关内容修改完成。

52．“模因”用到了文中列出的所有相关这个概念的出版物。

第六部分　关于意识的思考工具

53．“对立的两幅意象”中包括的素材取自*Consciousness Explained*（Dennett, 1991a）和*Sweet Dreams*（Dennett, 2005b）。

54．“僵尸直觉”，此概念是在*Sweet Dreams*（Dennett, 2005b）中被定义的。

55．“僵尸和殭尸”，根据*Consciousness Explained*（Dennett, 1991a）中的相关素材修改完成。

56．“花椰菜的诅咒”，根据“Quining Qualia”（Dennett, 1988a）中的相关素材修改完成。

57．“‘美元活力’究竟值多少钱”，根据“Consciousness: ‘How Much Is That inReal Money?’”（Dennett, 2001b）中的相关内容修改完成。

58．“卡普格拉先生的悲惨境况”，根据*Sweet Dreams*（Dennett 2005b）中的相关内容修改完成。

59．“听声辨牌”，根据“Explaining the ‘Magic’ of Consciousness”（Dennett, 2001c）中的相关内容修改完成。

60．“‘中文屋’错在哪里”，根据*The Mind's I*（Hofstadter and Dennett, 1981）和*The Intentional Stance*（Dennett, 1987）中的相关素材修改完成。

61．“从火星来的远程克隆人”，根据*The Mind's I*（Hofstadter and Dennett, 1981）中的相关内容修改完成。

62．“叙事重心”，这件思考工具最早出现于“Why Everyone Is a Novelist”（Dennett, 1988b）。

63．“他者现象学”，在“How to Study Consciousness Empirically: or Nothing Comes to Mind”（Dennett, 1982b）中首次提出，在*Consciousness Explained*（Dennett, 1991a）的一个章节中得到进一步阐述。由阿尔瓦·诺埃（Alva Noë）担任客座编辑发行的*Phenomenology and Cognitive Science*（vol. 6, nos. 1 and 2）特刊专门研究异现象学，有关这一主题最好的后续思考基本都包含其中。其中的“Heterophenomenology Reconsidered”（Dennett, 2007b）是我对这些思考的回复。

64．“色彩科学家玛丽：揭露吊杆托架”，根据*Consciousness Explained*（Dennett, 1991a）和“What RoboMary Knows”（Dennett, 2007d）中的相关内容修改完成。

第七部分　关于自由意志的思考工具

65．“一位恶毒的神经外科医生”出自我2012年的伊拉斯谟演讲“Erasmus: Sometimes a Spin-Doctor Is Right”。

66．“生命游戏”。马丁·加德纳（Martin Gardner）曾分别于1970年

10月和1971年2月，在《科学美国人》发表的两篇有关“数学游戏”的专栏文章中，向广大读者介绍了生命游戏。1985年，庞德斯通对生命游戏展开了一次出色探索并尝试揭露了其哲学内涵，其中还提到维基百科上一篇不错的相关文章，同时附带很多链接。我经常在写作中使用生命游戏。这次的版本是根据*Freedom Evolves*（Dennett, 2003）中的相关描述修订完成的。

67．“石头、剪子、布”之前从未发表过。

68．“两种彩票”，根据*Elbow Room*（Dennett, 1984a）中的相关内容修改完成。

69．“无效历史事实”是根据“Get Real”（Dennett, 1994a）修改完成的。Tanmoy Bhattacharya成功吸引了我对量子复杂性的关注。

70．“一场计算机国际象棋马拉松”出自*Freedom Evolves*（Dennett, 2003）。

71．“终极责任”之前从未发表过。

72．“掘土蜂状”之前从未发表过。

73．“巴西来的男孩”之前从未发表过。

第八部分　做个哲学家是一种怎样的体验

74．“浮士德式交易”是根据我针对托马斯·内格尔的书*Other Minds: Critical Essays, 1969–1994*写的一篇评论（Dennett, 1996b）修订整理的。

75．“简单的自我人类学”，根据“Sakes and Dints”（Dennett, 2012）和“Kinds of Things”（Dennett，即将发表）中的相关内容修改完成。

76．“‘像棋’中的二阶真理”，根据“The Higher-Order Truths about Chmess”（Dennett, 2006c）修改完成。

77．“只需关注那10%的精品”之前从未发表过。

参考文献

ABBOTT, EDWIN A., (1884) 1983, *Flatland: A Romance in Many Dimensions* (reprint of 1963 fifth edition with foreword by Isaac Asimov). New York: HarperCollins.

ALAIN [Émile-Auguste Chartier], (1908) 1956, *Propos d'un Normand 1906–1914*. Paris: Gallimard. Quoted in Deborah S. Rogers and Paul R. Ehrlich, 2008, "Natural Selection and Cultural Rates of Change." *Proceedings of the National Academy of Sciences*, vol. 105, pp.3416–3420.

AUSTIN, J. L., 1961, *Philosophical Papers*. Oxford: Oxford University Press.

AXELROD, ROBERT, 1984, *The Evolution of Cooperation.* New York: Basic Books.

AXELROD, ROBERT, and WILLIAM HAMILTON, 1981, "The Evolution of Cooperation." *Science*, vol. 211, pp. 1390–1396.

BARRETT, JUSTIN, 2000, "Exploring the Natural Foundations of Religion." *Trends in Cognitive Science*, vol. 4, pp. 29–34.

BATESON, WILLIAM, 1916, Review of the *Mechanisms of Mendelian Heredity* by T. H. Morgan (1914).

BEHE, MICHAEL J., 1996, *Darwin's Black Box: The Biochemical Challenge to Evolution*. New York: Free Press.

BENNETT, MAX, DANIEL DENNETT, P. M. S. HACKER, and JOHN SEARLE, 2009, *Neuroscience and Philosophy: Brain, Mind, and Language*. New York: Columbia University Press.

BENNETT, MAX, and P. M. S. HACKER, 2003, *Philosophical Foundations of Neuroscience*. Malden, Mass.: Wiley-Blackwell.

BLOCK, NED, 1978, "Troubles with Functionalism." In W. Savage, ed., *Perception and Cognition: Issues in the Foundations of Psychology*.

Minnesota Studies in the Philosophy of Science, vol. 9. Minneapolis: University of Minnesota Press, pp. 261–326.

BLOCK, NED, 1994, "What Is Dennett's Theory a Theory of?" *Philosophical Topics*, vol. 22 (special issue on the philosophy of Daniel Dennett), pp. 23–40.

BOBROW, DANIEL, 1985, *Qualitative Reasoning about Physical Systems*. Cambridge, Mass.: MIT Press.

BORGES, J. L., 1962, *Labyrinths: Selected Stories and Other Writings*. New York: New Directions.

BRAY, DENNIS, 2009, *Wetware*. New Haven, Conn.: Yale University Press.

BROCKMAN, J., ED., 2009, *What Have You Changed Your Mind About*. New York: HarperCollins.

BROCKMAN, J., ED., 2012, *This Will Make You Smarter*. New York: Harper Torchbook.

BROOKS, R. A., 1987, *Planning Is Just a Way of Avoiding Figuring Out What to Do Next*. Technical report, MIT Artificial Intelligence Laboratory, Cambridge, Mass. Available at http://people.csail.mit.edu/brooks/papers/Planning%20is%20Just.pdf.

BROOKS, RODNEY, 1991, "Intelligence without Representation." *Artificial Intelligence*, vol. 47, pp. 139–159.

CANN, REBECCA L., MARK STONEKING, and ALLAN C. WILSON, 1987, "Mitochondrial DNA and Human Evolution." *Nature*, vol. 325, pp. 31–36.

CHALMERS, DAVID, 1995, "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies*, vol. 2, no. 3, pp. 200–219.

CLARK, A., 2013, "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioral and Brain Sciences*.

CLEGG, LIAM, 2012, *Protean Free Will*. California Institute of

Technology, Pasadena.Available at http:// authors.library.caltech.edu/29887/.

COPE, DAVID, 2000, *The Algorithmic Composer.* Middleton, Wisc.: A-R Editions.

COPE, DAVID, 2001, *Virtual Music: Computer Synthesis of Musical Style*. Cambridge, Mass.: MIT Press.

CRONIN, HELENA, 1991, *The Ant and the Peacock.* Cambridge: Cambridge University Press.

DAMIAAN, DENYS, MARISKA MANTIONE, MARTIJN FIGEE, PEPIJN VAN DEN MUNCKHOF, FRANK KOERSELMAN, HERMAN WESTENBERG, ANDRIES BOSCH, and RICK SCHUURMAN, 2010, "Deep Brain Stimulation of the Nucleus Accumbens for Treatment-Refractory Obsessive-Compulsive Disorder." *Archives of General Psychiatry*, vol. 67, no. 10, pp. 1061–1068.

DAVIDSON, DONALD, 1987, "Knowing One's Own Mind." *Proceedings and Addresses of the American Philosophical Association*, vol. 60, pp. 441–458. Reprinted in Davidson, Donald, 2001, *Subjective, Intersubjective, Objective*. New York: Oxford University Press, pp.15–38.

DAWKINS, RICHARD, 1976, *The Selfish Gene*. Oxford: Oxford University Press. Rev.ed. 1989.

DAWKINS, RICHARD, 1982, *The Extended Phenotype: The Gene as the Unit of Selection*. Oxford: Oxford University Press.

DAWKINS, RICHARD, 1986, *The Blind Watchmaker*. London: Longmans.

DAWKINS, RICHARD, 1996, *Climbing Mount Improbable*. London: Viking Penguin.

DAWKINS, RICHARD, 2004, *The Ancestor's Tale: A Pilgrimage to the Dawn of Time*. London: Weidenfeld & Nicolson.

DE VRIES, PETER, 1953, *The Vale of Laughter*. Boston: Little, Brown.

DEHAENE, S., and J. F. MARQUES, 2002, "Cognitive Euroscience:

Scalar Variability in Price Estimation and the Cognitive Consequences of Switching to the Euro." *Quarterly Journal of Experimental Psychology*, vol. 55, pp. 705–731.

DENNETT, DANIEL C., 1968, "Features of Intentional Action." *Philosophy and Phenomenological Research*, vol. 29 (December), pp. 232–244.

DENNETT, DANIEL C., 1969, *Content and Consciousness*. London: Routledge & Kegan Paul.

DENNETT, DANIEL C., 1971, "Intentional Systems." *Journal of Philosophy*, vol. 68, pp. 87–106.

DENNETT, DANIEL C., 1975, "Brain Writing and Mind Reading." In K. Gunderson, ed., *Language, Mind and Knowledge.* Minnesota Studies in the Philosophy of Science, vol. 7. Minneapolis: University of Minnesota Press, pp. 403–416. Reprinted in Dennett 1978a.

DENNETT, DANIEL C., 1978a, *Brainstorms*. Cambridge, Mass.: MIT Press/A Bradford Book.

DENNETT, DANIEL C., 1978b, "Current Issues in the Philosophy of Mind." *American Philosophical Quarterly*, vol. 15, pp. 249–261.

DENNETT, DANIEL C., 1978c, "Why You Can't Make a Computer That Feels Pain." *Synthese*, vol. 38 (August), pp. 415–456.

DENNETT, DANIEL C., 1980, "The Milk of Human Intentionality" (commentary on Searle). *Behavioral and Brain Sciences*, vol. 3, pp. 428–430.

DENNETT, DANIEL C., 1981, "Three Kinds of Intentional Psychology." In R. Healey, ed., *Reduction, Time and Reality*. Cambridge: Cambridge University Press, pp. 37–61.

DENNETT, DANIEL C., 1982a, "Beyond Belief." In A. Woodfield, ed., *Thought and Object: Essays on Intentionality*. Oxford: Oxford University Press. Reprinted in Dennett, 1987.

DENNETT, DANIEL C., 1982b, "How to Study Consciousness Empirically: or Nothing Comes to Mind." *Synthese*, vol. 53, pp. 159–180.

DENNETT, DANIEL C., 1982c, "The Myth of the Computer: An Exchange" (reply to John Searle's review of *The Mind's I*). *New York Review of Books*, vol. 29 (June 24), pp. 56–57.

DENNETT, DANIEL C., 1982d, "Notes on Prosthetic Imagination." *New Boston Review*, vol. 7 (June), pp. 3–7. Reprinted in "30 Years of Boston Review." *Boston Review*, vol. 30, no.5 (September/October 2005), p. 40.

DENNETT, DANIEL C., 1983, "Intentional Systems in Cognitive Ethology: The 'Panglossian Paradigm' Defended." *Behavioral and Brain Sciences*, vol. 6, pp. 343–390.

DENNETT, DANIEL C., 1984a, *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, Mass.: MIT Press.

DENNETT, DANIEL C., 1984b, "I Could Not Have Done Otherwise —So What?" *Journal of Philosophy*, vol. 81, pp. 553–565.

DENNETT, DANIEL C., 1986, "The Logical Geography of Computational Approaches: A View from the East Pole." In Robert M. Harnish and M. Brand, eds., *The Representation of Knowledge and Belief*. Tucson: University of Arizona Press, pp. 59–79.

DENNETT, DANIEL C., 1987, *The Intentional Stance*. Cambridge, Mass.: MIT Press.

DENNETT, DANIEL C., 1988a, "Quining Qualia." In A. Marcel and E. Bisiach, eds., *Consciousness in Modern Science.* Oxford: Oxford University Press, pp. 42–77.

DENNETT, DANIEL C., 1988b, "Why Everyone Is a Novelist." *Times Literary Supplement*, vol. 4 (September 16–22), p. 459.

DENNETT, DANIEL C., 1988c, "Evolution, Error and Intentionality." In Y. Wilks and D. Partridge, eds., *Sourcebook on the Foundations of Artificial Intelligence*. Albuquerque: University of New Mexico Press, pp. 190–211.

DENNETT, DANIEL C., 1990, "Memes and the Exploitation of Imagination." *Journal of Aesthetics and Art Criticism*, vol. 48, pp. 127–135.

DENNETT, DANIEL C., 1991a, *Consciousness Explained*. Boston: Little, Brown.

DENNETT, DANIEL C., 1991b, "Real Patterns." *Journal of Philosophy*, vol. 88, pp.27–51.

DENNETT, DANIEL C., 1993, "Confusion over Evolution: An Exchange." *New York Review of Books*, January 14, pp. 43–44.

DENNETT, DANIEL C., 1994a, "Get Real" (reply to my critics). *Philosophical Topics*, vol. 22 (special issue on the philosophy of Daniel Dennett), pp. 505–556.

DENNETT, DANIEL C., 1994b, "The Practical Requirements for Making a Conscious Robot." *Proceedings of the Royal Society*, *A*, vol. 349, pp. 133–146.

DENNETT, DANIEL C., 1995a, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.

DENNETT, DANIEL C., 1995b, "How to Make Mistakes." In J. Brockman and K. Matson, eds., *How Things Are*. New York: William Morrow, pp. 137–144.

DENNETT, DANIEL C., 1996a, *Kinds of Minds: Towards an Understanding of Consciousness.* New York: Basic Books.

DENNETT, DANIEL C., 1996b, "Review of *Other Minds: Critical Essays, 1969–1994* by Thomas Nagel, 1995." *Journal of Philosophy*, vol. 63, no. 8 (August), pp. 425–428.

DENNETT, DANIEL C., 2000, "With a Little Help from My Friends." In Don Ross, Andrew Brook, and David Thompson, eds., *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, Mass.: MIT Press, pp. 327–388.

DENNETT, DANIEL C., 2001a, "Collision-Detection, Muselot, and Scribble: Some Reflections on Creativity." In Cope, 2001, pp. 283–291.

DENNETT, DANIEL C., 2001b, "Consciousness: How Much Is That in Real Money?" In R. Gregory, ed., *The Oxford Companion to the Mind*, 2nd ed. Oxford: Oxford University Press.

DENNETT, DANIEL C., 2001c, "Explaining the 'Magic' of Consciousness." In *Exploring Consciousness, Humanities, Natural Science, Religion, Proceedings of the International Symposium, Milano, November 19-20, 2001* (published in December 2002, Fondazione Carlo Erba), pp. 47–58.

DENNETT, DANIEL C., 2001d, "In Darwin's Wake, Where Am I?" (American Philosophical Association Presidential Address). *Proceedings and Addresses of the American Philosophical Association*, vol. 75, no. 2 (November), pp. 13–30. Reprinted in J. Hodge and G. Radick, eds., 2003, *The Cambridge Companion to Darwin*. Cambridge: Cambridge University Press, pp. 357–376.

DENNETT, DANIEL C., 2001e, "Things about Things." In Joao Branquinho, ed., *The Foundations of Cognitive Science*. Oxford: Clarendon Press, pp. 133–149.

DENNETT, DANIEL C., 2002, "The New Replicators." In Mark Pagel, ed., *The Encyclopedia of Evolution*, vol. 1. Oxford: Oxford University Press, pp. E83–E92.

DENNETT, DANIEL C., 2003, *Freedom Evolves*. New York: Viking Penguin.

DENNETT, DANIEL C., 2004, "Holding a Mirror up to Dupré" (commentary on John Dupré, *Human Nature and the Limits of Science*). *Philosophy and Phenomenological Research*, vol. 69, no. 2 (September), pp. 473–483.

DENNETT, DANIEL C., 2005a, "Natural Freedom." *Metaphilosophy*, vol. 36, no. 4 (July), pp. 449–459.

DENNETT, DANIEL C., 2005b, *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness.* Cambridge, Mass.: MIT Press.

DENNETT, DANIEL C., 2006a, *Breaking the Spell: Religion as a Natural Phenomenon*. New York: Viking Penguin.

DENNETT, DANIEL C., 2006b, "From Typo to Thinko: When Evolution Graduated to Semantic Norms." In S. Levinson and P. Jaisson, eds., *Evolution and Culture*. Cambridge, Mass.: MIT Press, pp.

133–145.

DENNETT, DANIEL C., 2006c, "The Higher-Order Truths about Chmess." *Topoi*, pp.39–41.

DENNETT, DANIEL C., 2007a, "The God Delusion by Richard Dawkins." *Free Inquiry*, vol. 27, no. 1 (December/January 2007).

DENNETT, DANIEL C., 2007b, "Heterophenomenology Reconsidered." *Phenomenology and Cognitive Science*, vol. 6, nos. 1 and 2 (special issue on heterophenomenology, Alva Noë, ed.), pp. 247–270.

DENNETT, DANIEL C., 2007c, "Varieties of Content." Presentation at Concepts: Content and Constitution, A Symposium, University of Copenhagen, Amager, Denmark, May 12.

DENNETT, DANIEL C., 2007d, "What RoboMary Knows." In T. Alter and S. Walter, eds., *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press, pp. 15–31.

DENNETT, DANIEL C., 2008, "Competition in the Brain." World Question Center. Edge.org, December, later published in Brockman, 2009.

DENNETT, DANIEL C., 2009a, "Darwin's 'Strange Inversion of Reasoning.'" *Proceedings of the National Academy of the Sciences of the United States of America*, vol. 106, suppl. 1, pp. 10061–10065.

DENNETT, DANIEL C., 2009b, "Heterophenomenology." In T. Bayne, A. Cleeremans, and P. Wilken, eds., *The Oxford Companion to Consciousness*. Oxford: Oxford University Press, pp. 345–346.

DENNETT, DANIEL C., 2009c, "Intentional Systems Theory." In B. McLaughlin, A. Beckermann, and S. Walter, eds., *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press, pp. 339–350.

DENNETT, DANIEL C., 2010, "Homunculi Rule: *Reflections on Darwinian Populations and Natural Selection* by Peter Godfrey-Smith, Oxford University Press, 2009." *Biology and Philosophy* (published online December 21). Available at http://ase.tufts.edu/cogstud/

papers/homunculi.pdf.

DENNETT, DANIEL C., 2011a, "Cycles" (as an answer to the Edge Question 2011, "What Scientific Concept Would Improve Everybody's Cognitive Toolkit?"). World Question Center. Edge.org, later published in Brockman, 2012, pp. 81–88.

DENNETT, DANIEL C., 2011b, "Shall We Tango? No, but Thanks for Asking" (commentaries on Evan Thompson, *Mind in Life*, with replies). *Journal of Consciousness Studies*, vol. 18, nos. 5 and 6 (special issue on Evan Thompson), pp. 23–34.

DENNETT, DANIEL C., 2012, "Sakes and Dints." *Times Literary Supplement,* March 2, pp. 12–14.

DENNETT, DANIEL C., forthcoming, "The Evolution of Reasons." In BANA BASHOUR and HANS D. MULLER, eds., Contemporary Philosophical Naturalism and Its Implications. New York: Routledge.

DENNETT, DANIEL C., forthcoming, "Expecting Ourselves to Expect" (commentary on Clark). *Behavioral and Brain Sciences*.

DENNETT, DANIEL C., forthcoming, "Kinds of Things." In Don Ross, James Ladyman, and Harold Kincaid, eds., *Does Scientific Philosophy Exclude Metaphysics?* Oxford: Oxford University Press.

DENNETT, DANIEL C., forthcoming, "Turing's 'Strange Inversion of Reasoning.'" In Barry Cooper, ed., *Alan Turing—His Work and Impact.* Elsevier.

DENNETT, DANIEL C., and KATHLEEN AKINS, 2008, "The Multiple Drafts Model." *Scholarpedia,* vol. 3, no. 4, 4321. http://www.scholarpedia.org/wiki/index.php?title=Multiple_drafts_model.

DENNETT, DANIEL C., and ALVIN PLANTINGA, 2011, *Science and Religion: Are They Compatible?* Oxford: Oxford University Press.

DENNETT, DANIEL C., and C. F. WESTBURY, 2000, "Mining the Past to Construct the Future: Memory and Belief as Forms of Knowledge." In D. Schacter and E. Scarry, eds., *Memory, Brain, and Belief.* Cambridge, Mass.: Harvard University Press, pp. 11–32.

DENTON, MICHAEL, 1985, *Evolution: A Theory in Crisis*. London:

Burnett Books.

DEWDNEY, A. K., 1984, *The Planiverse: Computer Contact with a Two-Dimensional World*. New York: Poseidon Press.

EIGEN, MANFRED, 1992, *Steps towards Life*. Oxford: Oxford University Press.

ELLIS, HAYDN, and ANDREW YOUNG, 1990, "Accounting for Delusional Misidentifications." *British Journal of Psychiatry*, vol. 157, pp. 239–248.

FEYNMAN, RICHARD, 1985, *"Surely You're Joking, Mr. Feynman!": Adventures of a Curious Character*. New York: W. W. Norton.

FITCH, TECUMSEH, 2008, "Nano-Intentionality: A Defense of Intrinsic Intentionality." *Biology and Philosophy*, vol. 23, pp. 157–177.

FITZGIBBON, C. D., and J. H. FANSHAWE, 1988, "Stotting in Thomson's Gazelles: An Honest Signal of Condition." *Behavioral Ecology and Sociobiology*, vol. 23, no. 2 (August), pp. 69–74.

FODOR, JERRY, 1975, *The Language of Thought*. Hassocks, Sussex: Harvester Press.

FODOR, JERRY, 2003, "Why Would Mother Nature Bother?" (review of *Freedom Evolves*). *London Review of Books*, vol. 25, no. 5, pp. 17–18.

FODOR, JERRY, 2008, *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

FODOR, JERRY, and E. LEPORE, 1992, *Holism: A Shopper's Guide*. Oxford: Blackwell.

FODOR, JERRY, and M. PIATELLI-PALMERINI, 2010, *What Darwin Got Wrong*. New York: Farrar, Straus & Giroux.

FRANKFURT, HARRY, 1969, "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy*, vol. 65, pp. 829–833.

GELL-MAN, MURRAY, 1995, *The Quark and the Jaguar: Adventures in the Simple and the Complex*. New York: St. Martin's.

GIBSON, J. J., 1979, *The Ecological Approach to Visual Perception*.

Boston: Houghton Mifflin.

GODFREY-SMITH, PETER, 2009, *Darwinian Populations and Natural Selection*.Oxford: Oxford University Press.

GODFREY-SMITH, PETER, 2011, "Agents and Acacias: Replies to Dennett, Sterelny, and Queller" (replies to reviews of *Darwinian Populations and Natural Selection*). *Biology and Philosophy*, vol. 26, pp. 501–515.

GOFFMAN, ERVING, 1959, *The Presentation of Self in Everyday Life*. Edinburgh: University of Edinburgh Social Sciences Research Centre.

GOULD, STEPHEN JAY, 1977, *Ever since Darwin*. New York: W. W. Norton.

GOULD, STEPHEN JAY, 1989a, "Tires to Sandals." *Natural History*, April, pp. 8-15.

GOULD, STEPHEN JAY, 1989b, *Wonderful Life: The Burgess Shale and the Nature of History*. New York: W. W. Norton.

GOULD, STEPHEN JAY, 1992a, "The Confusion over Evolution." *New York Review of Books*, November 19.

GOULD, STEPHEN JAY, 1992b, "Life in a Punctuation." *Natural History*, October, pp.10–21.

GOULD, STEPHEN JAY, 1993, "Confusion over Evolution: An Exchange." *New York Review of Books*, January 14, pp. 43–44.

GRAHAM, GEORGE, and TERENCE HORGAN, 2000, "Mary Mary Quite Contrary." *Philosophical Studies*, vol. 99, pp. 59–87.

GREENE, JOSHUA, and JONATHAN COHEN, 2004, "For the Law, Neuroscience Changes Everything and Nothing." *Philosophical Transactions of the Royal Society*, vol. 359, pp. 1775–1785.

HARRIS, SAM, 2012, *Free Will*. New York: Free Press.

HAUGELAND, JOHN, 1981, *Mind Design*. Cambridge, Mass.: MIT Press/Bradford Book.

HAUGELAND, JOHN, 1985, *Artificial Intelligence: The Very Idea*.

Cambridge, Mass.:MIT Press.

HAWKING, STEPHEN W., 1988, *A Brief History of Time*. New York: Bantam.

HAYES, PATRICK, 1978, "The Naïve Physics Manifesto." In D. Michie, ed., *Expert Systems in the Microelectronic Age*. Edinburgh: Edinburgh University Press.

HEIDER, F., and M. SIMMEL, 1944, "An Experimental Study of Apparent Behavior." *American Journal of Psychology*, vol. 57, no. 2, pp. 243–259.

HILLIARD, JOHN NORTHERN, 1938, *Card Magic*. Minneapolis: Carl W. Jones.

HOFSTADTER, DOUGLAS, 1979, *Gödel Escher Bach*. New York: Basic Books.

HOFSTADTER, DOUGLAS, 1982, "Metamagical Themas: Can Inspiration Be Mechanized?" *Scientific American*, September, pp. 18–34. Reprinted as "On the Seeming Paradox of Mechanizing Creativity," in Hofstadter, 1985, pp. 526–546.

HOFSTADTER, DOUGLAS, 1985, *Metamagical Themas: Questing for the Essence of Mind and Pattern*. New York: Basic Books.

HOFSTADTER, DOUGLAS, 1997, *Le Ton Beau de Marot: In Praise of the Music of Language*. New York: Basic Books.

HOFSTADTER, DOUGLAS, 2007, *I Am a Strange Loop*. New York: Basic Books.

HOFSTADTER, DOUGLAS, and DANIEL DENNETT, eds., 1981, *The Mind's I*. New York: Basic Books.

HOLLAND, JOHN, 1975, *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

HUME, DAVID, (1739) 1964, *A Treatise of Human Nature* (L. A. Selby-Bigge, ed.). Oxford: Clarendon.

HUMPHREY, NICHOLAS, 1987, "Scientific Shakespeare." *The Guardian* (London), August 26.

HUMPHREY, NICHOLAS, and DANIEL DENNETT, 1989, "Speaking for Our Selves: An Assessment of Multiple Personality Disorder." *Raritan: A Quarterly Review*, vol. 9(Summer), pp. 68–98. Reprinted (with footnotes), Occasional Paper 8, Center on Violence and Human Survival, John Jay College of Criminal Justice, City University of New York, 1991.

HURLEY, MATTHEW, DANIEL DENNETT, and REGINALD B. ADAMS JR., 2011, *Inside Jokes: Using Humor to Reverse-Engineer the Mind*. Cambridge, Mass.: MIT Press.

JACKENDOFF, RAY, 1987, *Consciousness and the Computational Mind*. Cambridge, Mass.: MIT Press/Bradford Book.

JACKENDOFF, RAY, 1993, *Patterns in the Mind: Language and Human Nature*.Harlow, Essex: Harvester Wheatsheaf; New York: Basic Books, 1994.

JACKSON, FRANK, 1982, "Epiphenomenal Qualia." *Philosophical Quarterly*, vol. 32, pp. 127–136.

KANE, ROBERT, 1996, *The Significance of Free Will*. Oxford: Oxford University Press.

KNOBE, J., and NICHOLS, S., eds., 2008, *Experimental Philosophy*. Oxford: Oxford University Press.

LEIBNIZ, GOTTFRIED, (1714) 1898, *Monadology*. In *The Monadology and Other Philosophical Writings*, Robert Latta, trans. Oxford: Oxford University Press.

LENAT, DOUGLAS B., and R. V. GUHA, 1990, *Building Large Knowledge-Based Systems: Representation and Inference in the CYC Project*. Reading, Mass.: Addison-Wesley.

LETTVIN, J. Y., U. MATURANA, W. MCCULLOCH, and W. PITTS, 1959, "What the Frog's Eye Tells the Frog's Brain." In *Proceedings of the Institute of Radio Engineers*, vol. 47, pp. 1940–1951.

LEVINE, JOSEPH, 1994, "Out of the Closet: A Qualophile Confronts Qualophobia." *Philosophical Topics*, vol. 22, pp. 107–126.

LEWIS, DAVID, 1978, "Truth in Fiction." *American Philosophical*

Quarterly, vol. 15, pp. 37–46.

LLOYD, M., and DYBAS, H. S., 1966, "The Periodical Cicada Problem." *Evolution*, vol.20, pp. 132–149.

LLOYD MORGAN, CONWY, 1894, *An Introduction to Comparative Psychology.*London: W. Scott.

LUDLOW, PETER, YUJIN NAGASAWA, and DANIEL STOLJAR, eds., 2004, *There's Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. Cambridge, Mass.: MIT Press/Bradford Books.

MACKENZIE R. B., 1868, *The Darwinian Theory of the Transmutation of Species Examined.* London: Nisbet.

MAYNARD SMITH, JOHN, 1978, *The Evolution of Sex*. Cambridge: Cambridge University Press.

MCCLELLAND, JAY, DAVID RUMELHART, and THE PDP RESEARCH GROUP, 1986, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 2. Cambridge, Mass.: MIT Press.

MENABREA, LUIGI FEDERICO, 1842, "Sketch of the Analytic Engine Invented by Charles Babbage." In the *Bibliothèque Universelle de Genève*, no. 82 (October). Translated by Augusta Ada King, Countess of Lovelace, 1843, with notes, in *Scientific Memoirs*, vol. 3, pp.666–731.

MEYER, STEPHEN C., 2009, *Signature in the Cell: DNA and the Evidence for Intelligent Design*. New York: HarperOne.

MILLIKAN, RUTH, 1984, *Language, Thought and Other Biological Categories*. Cambridge, Mass.: MIT Press.

MILLIKAN, RUTH, 1993, *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass.: MIT Press.

NAGEL, THOMAS, 1974, "What Is It Like to Be a Bat?" *Philosophical Review,* vol. 83, pp. 435–450.

NAGEL, THOMAS, 2009, Recommendation for Book of the Year. *Times Literary Supplement*, November 27.

NAGEL, THOMAS, 2010, Letter to the editor. *Times Literary*

Supplement, January 1.

NEUGEBAUER, OTTO, 1989, "A Babylonian Lunar Ephemeris from Roman Egypt." In E. Leichty, M. de J. Ellis, and P. Gerardi, eds., *A Scientific Humanist: Studies in Honor of Abraham Sachs*. Philadelphia: Occasional Publications of the Samuel Noah Kramer Fund no. 9, pp. 301–304.

PINKER, STEVEN, 2002, *The Blank Slate: The Modern Denial of Human Nature*. New York: Viking.

POPEK, GERALD J., and ROBERT P. GOLDBERG, 1974, "Formal Requirements for Virtualizable Third Generation Architectures." *Communications of the ACM*, vol. 17, no. 7, pp.412–421. doi:10.1145/361011.361073. Available at http://doi.acm.org/10.1145/361011.361073.

POUNDSTONE, WILLIAM, 1985, *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*. New York: William Morrow.

PUTNAM, HILARY, 1975, "The Meaning of 'Meaning.'" In K. Gunderson, ed., *Language, Mind and Knowledge*. Minnesota Studies in the Philosophy of Science, vol. 7. Minneapolis: University of Minnesota Press. Reprinted in Putnam, 1975, *Mind, Language and Reality (Philosophical Papers*, vol. 2). Cambridge: Cambridge University Press.

QUINE, W. V. O., 1960, *Word and Object*. Cambridge, Mass.: MIT Press.

QUINE, W. V. O., 1987, "Universal Library." In *Quiddities: An Intermittently Philosophical Dictionary*. Cambridge, Mass.: Harvard University Press.

RAPOPORT, ANATOL, 1960, *Fights, Games, and Debates*. Ann Arbor: University of Michigan Press.

RAPOPORT, ANATOL, 1961, "Three Modes of Conflict." *Management Science*, vol. 3, p.210.

RIDLEY, MATT, 1993, *The Red Queen: Sex and the Evolution of*

Human Nature. New York: Macmillan.

RIDLEY, MATT, 2004, *Nature via Nurture*. London: Fourth Estate. Also published under the title *The Agile Gene: How Nature Turns on Nurture*. New York: HarperCollins.

ROSS, AMBER, 2013, "Inconceivable Minds." Philosophy PhD dissertation, University of North Carolina at Chapel Hill.

RUINA, ANDY, 2011, "Cornell Ranger, 2011, 4-Legged Bipedal Robot." ruina.tam.cornell.edu/research/topics/locomotion_and_robotics/ranger/Ranger 2011/.

RUMELHART, D. E., J. L. McCLELLAND, and THE PDP RESEARCH GROUP, 1986, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1. Cambridge, Mass.: MIT Press.

RYDER, DAN, JUSTINE KINGSBURY, and KENNETH WILLIFORD, eds., 2013, *Millikan and Her Critics*. Oxford: Wiley-Blackwell.

SANFORD, DAVID, 1975, "Infinity and Vagueness." *Philosophical Review*, vol. 84, pp.520–535.

SCHÖNBORN, CHRISTOPH, 2005, "Finding Design in Nature." *New York Times*, July 7.

SEARLE, JOHN, 1980, "Minds, Brains and Programs." *Behavioral and Brain Sciences*, vol. 3, pp. 417–458.

SEARLE, JOHN, 1982, "The Myth of the Computer" (review of *The Mind's I*). *New York Review of Books*, vol. 29 (April 29).

SEARLE, JOHN, 1988, "Turing the Chinese Room." In T. Singh, ed., *Synthesis of Science and Religion, Critical Essays and Dialogues*. San Francisco: Bhaktivedanta Institute.

SELLARS, WILFRID, 1962, "Philosophy and the Scientific Image of Man." In *Science, Perception and Reality*. London: Routledge & Kegan Paul.

SEUNG, SEBASTIAN, 2007, "The Once and Future Science of Neural Networks." Presented at the Society for Neuroscience meeting, San Diego, November 4.

SIEGEL, LEE, 1991, *Net of Magic: Wonders and Deceptions in India*. Chicago: University of Chicago Press.

SIMS, KARL, 1994, *Evolved Virtual Creatures*. http://www.karlsims.com/evolved-virtual-creatures.html.

STRAWSON, GALEN, 2003, "Evolution Explains It All for You" (review of Dennett, 2003). *New York Times*, March 2.

STRAWSON, GALEN, 2010, "Your Move: The Maze of Free Will," The Stone, *New York Times* online, July 22, 2010. http://www.scribd.com/doc/86763712/Week-2-Strawson-The-Maze-of-Free-Will.

THOMPSON, EVAN, 2007, *Mind in Life*. Cambridge, Mass.: Belknap Press, Harvard University Press.

VOHS, KATHLEEN D., and JONATHAN W. SCHOOLER, 2008, "The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating." *Psychological Science*, pp. 49–54.

VON NEUMANN, JOHN, 1966, *Theory of Self-Reproducing Automata* (Arthur Burks, ed.). Champaign-Urbana: University of Illinois Press.

VON NEUMANN, JOHN, and OSKAR MORGENSTERN, 1944, *Theory of Games and Economic Behavior.* Princeton, N.J.: Princeton University Press.

VON UEXKÜLL, JAKOB, (1934) 1957, "A Stroll through the Worlds of Animals and Men: A Picture Book of Invisible Worlds." In Claire H. Schiller, ed. and trans., *Instinctive Behavior: The Development of a Modern Concept*. New York: International Universities Press.

WANG, HAO, 1957, "A Variant to Turing's Theory of Computing Machines." *Journal of the Association for Computing Machinery*, pp. 63–92.

WIGGINS, DAVID, 1973, "Towards a Reasonable Libertarianism." In T. Honderich, ed., *Essays on Freedom of Action*. London: Routledge & Kegan Paul, pp. 31–63.

WIMSATT, WILLIAM C., 1980, "Randomness and Perceived

Randomness in Evolutionary Biology." *Synthese*, vol. 43, pp. 287–290.

WOLFE, TOM, 2000, "Sorry, But Your Soul Just Died." In *Hooking Up*. New York: Farrar, Straus & Giroux.

WOOLDRIDGE, DEAN, 1963, *The Machinery of the Brain*. New York: McGraw-Hill.

WRIGHT, ROBERT, 2000, *Nonzero: The Logic of Human Destiny*. New York: Pantheon.

ZAHAVI, A., 1987, "The Theory of Signal Selection and Some of Its Implications." In V.P. Delfino, ed., *Bari, 9-14 April 1985*. Bari: Adriatici Editrici, pp. 305–327.

译者后记

记得刚入学攻读哲学硕士、正经开始接触西方哲学的时候，专业课老师就常常教导我们：国内的西方哲学学习，翻译原著文本是必需功课。但众多师兄师姐会拿经验告诉你，真正展开一项翻译工作其实大多都有些“费力不讨好”，因为它非常耗时耗力，却又不能真正算作自己的学术成果。所幸《直觉泵》一书的整个翻译过程虽然艰辛漫长，但并不苦闷。丹尼尔对诸多问题的思考亮点频繁闪现、充满原创性，让我们每天通过译读能感受到一些实在的收获。而对于一些掌舵方向的“大问题”，比如在当下哲学环境中，如何正确看待科学与哲学的关系，如何做哲学、学习哲学等等，虽然文中没有作为专门问题提出，但内容阐述中透露出来的那些成熟稳健的相关思考，即便对于已经学习和研究过西方哲学多年的学生甚至学者，也都是有所助益的。

相比于其他还算按部就班的哲学作品，丹尼特此书的风格其实可算是一次新尝试。就像他在前言中所说，本书要写成一本人人都能读懂的哲学书，所以在他一贯的初稿试讲测试中，他排除掉了讨论班中所有哲学专业的学生和研究生，只留下十几个没有太多知识背景的本科新生，他们可以对书中任何不懂的细节提出疑问，并不断要求丹尼特给出清晰通俗的解释。可能正是因为这个原因，《直觉泵》中能看到很多有意思的设想、疑问和隐喻，很多细节问题的剖析也都向外伸展得非常细长。读者在阅读过程中，可能会因为能跟随一位大思想家一步步完成每一次漂亮的论证和推理而感到欣喜，但也难免会因为脱离其思想主干太远、不明所以而生出苦恼。

毕竟我们都能看到，丹尼特不是真的在写一本“工具书”，而事实上，本书内容基本涵盖当代心灵哲学研究所有关键的概念和论题，包括意识、意向性、意义、自我、自由意志等等，所以，稍微对哲学有点关心的读者都更渴望能直接看到丹尼特坚持的观点和立场。但这种“看似的偏离主题”其实并不真正构成什么舛误，相反，更专业些的学者们会告诉你，那些看似拓展说明的细节、案例或者假设其实在心灵哲学的大背景中都有所针对，所以自然携带着作者的主导观点和立场信息。另外，丹尼特在进化论基础上坚持自然主义的心灵理论，由进化论铺展开来的对客观存在与意义世界的统一解释几乎就是自然主义说明的一种标准进

路。所以，那种囊括各种可能性、无限伸展的表达方式实则也是丹尼特心灵哲学在进化论指导下阐明心灵状态的一种展现。

感谢我的两位老师，陈嘉映老师和叶峰老师。翻译过程中当然会遇到困难，但让人更苦恼的是，时不时要面临赤裸裸的智力挑衅，《直觉泵》中太多反直观的优质论证让人不得不承认，大脑与大脑间可能真的存在智力差距。做陈老师的学生到2018年这个月正好整十年，学生当得久了，心情轻松些了，自然是要把这些懊恼说与自己的老师听听。而陈老师的一句“只要熟悉了还是可以得心应手”也让我如愿瞬间重拾自信、脱离苦海。陈老师总是这样，不经意间就慷慨分发我们些小小“法宝”，估计如果使用得当，再用个十几二十年也不见得会显老旧。

其实在邀请叶老师为此书写序之前，我与叶老师并不十分熟悉，因为分属不同的学科方向，几年时间的博士学习期间，我甚至没有系统选修过老师任何一门课程。所以我猜想，那天叶老师在接到来自学生邀请的这份“苦差事”时大概都有些惊异，才会弱弱问我，如果他来写序能给我们带来什么实质性的帮助？我顿时有感而发，把这部著作的表述特点和可能给读者带来的一些混淆统统说出，没想到叶老师听后便立马答应下来。不久后，收到老师一篇非常出色、近万字的导论性序言。感激、感动。

还要感谢我们的编辑老师简学、赵地和王子涵，不管是在翻译阶段还是在修改阶段，几位编辑对我们的纠结、拖沓、犹豫和举棋不定都表现出了超乎常人的耐心和宽容。金岳与我是嘉映老师在首都师范大学招收的第一批硕士研究生，有过一年上无师兄师姐、下无师弟师妹的独特时光，同门情谊格外深厚；徐韬天生性情稳重、聪明敏锐，中英文水平俱佳，也有一些西方哲学辑写和翻译的经验，是我俩的首位直系师弟，更是不可多得的合作伙伴。整个翻译过程我们配合默契、有商有量，非常愉快顺畅。此书在翻译过程中尽量保留了丹尼特原本行文中轻松流畅的表述风格，希望读者们喜欢；对于文中不足，请多多包涵，也欢迎大家的指导和建议。

冯文婧
2018年8月

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读APP：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”APP。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读APP玩转指南

湛庐阅读APP结构图：



三步玩转湛庐阅读APP：



APP获取方式：
安卓用户前往各大应用市场、苹果用户前往APP Store
直接下载"湛庐阅读"APP，与最聪明的人共同进化！

使用APP扫一扫功能，遇见书里书外更大的世界！



快速了解本书内容，
湛庐千册图书一键购买！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

延伸阅读

《心智探奇》

◎ 当代伟大思想家、TED演讲人、世界顶级语言学家和认知心理学家史蒂芬·平克经典重现。

◎ 认知神经科学领域颠覆性著作，凝聚认知神经学、人工智能和进化心理学等多项研究成果。

◎ 观点独到精辟，立论严谨周密，论证新颖犀利，例证丰富新鲜，行文汪洋恣肆。





《白板》

◎ 史蒂芬·平克“语言与人性”四部曲之四，继《语言本能》《思想本质》《心智探奇》后又一力作。

◎ 关于主流人性论的颠覆性反叛，破而后立，进而建立起新的意义和道德观念，重塑对人性的信心。

◎ 阐明人性的本源、内涵及局限，追溯人类诸多苦难的根源，拨开道德错觉的迷雾，直达现实的彼岸。





《笛卡尔的错误》

◎ 国内外众多知名神经科学家及诺贝尔奖得主集体盛赞，北京大学国家发展研究院经济学教授汪丁丁倾情作序。

◎ 知名神经科学家达马西奥颠覆性巨著，标志着20世纪时代思想的转折，神经科学、心理学和哲学研究因此出现了巨大的转向。

◎ 一场人类意识探索的盛宴，一次刷新阅读体验的脑力挑战。犹如悬疑小说一般娴熟的叙事技巧，在哲学和科学的交界处跳着踢踏舞。





《简单的哲学》

◎ 清华大学、中国人民大学、南京师范大学，3大高校哲学教授倾情推荐。

◎ 英国《哲学家杂志》主编、特兰西瓦尼亚大学杰出教授联袂奉献。

◎ 《你以为你以为的就是你以为的吗》作者经典著作。



使用"湛庐阅读"APP，
"扫一扫"获取本书更多精彩内容



(1) 《费曼物理学讲义》第一卷，理查德·费曼著。

(2) 《谈谈物理学研究和教学》，是杨振宁1986年在中国科学技术大学研究生院的谈话。

(3) 伊索和荷马很像，都同自己的作品一样神秘。他的寓言在口头流传了数个世纪之后才被记录，这比柏拉图和苏格拉底时代还要早几百年。伊索很可能不是希腊人，有间接证据表明他是埃塞俄比亚人。

(4) 侯世达的新作《表象与本质》已由湛庐文化引进，即将于2018年出版。——编者注

(5) 哲学家W. V. O.蒯因（W. V. O. Quine）称其为语义上行（semantic ascent），我们从谈论电子、正义、马等事物，上行到谈论电子、正义、马等词语的语义。有时人们反对哲学家们这么做："恕我直言，你搞的全是语义学而已！"有时这么做确实没有用，甚至会产生误导，但当我们需要这么做时，就是当我们开始彼此自说自话，被所用词语中的一些心照不宣的假设所蒙蔽时，语义上行或者"溯元"，就成了我们澄清问题的关键。

(6) 这本书中的很多段落出自我以前出版过的书和文章，它们被我改动得更加灵活、通用，像大多数好的思考工具一样，可以适用于各种各样的上下文中。比如，开篇所提到的冯·诺依曼的故事出自我1995年写的《达尔文的危险观念》（*Darwin's Dangerous Idea*），而我于2009年发表在《美国科学院院报》（*PNAS*）中的文章《达尔文"奇怪的推理倒置"》（*Darwin's "Strange Inversion of Reasoning"*）中讨论过侯世达列举的这些思考工具。我会在书后列一个表，统一标明出处，所以就不再一一注释了。

(7) 这两句挖苦得最妙："哲学之于科学，就像广场上的鸽子之于雕像。""哲学之于科学，就像色情书刊之于性：它更便宜、更容易获得，有些人也更喜欢它。"（在这里我就不点名了，不过如果作者愿意，他们可以站出来承认。）

(8) 水手们并不能一下子确定他们所处的准确位置，因为一次计算并不能确定到一个点，但是可以确定到一条线。你在这条位置线的某处等待几小时，等待太阳明显移动一定距离之后，在你已经获得的位置线上任取一点，假设此点就是你所处的准确位置，并算出太阳应处的位置。然后再观察太阳的实际位置，比较两者，做出修正，就得到另外一条位置线。前后两次得出的两条位置线的交点就是你实际所处的位置了。在你等待的几个小时中，太阳不仅会改变高度，还会改变方位，所以两条线会以某个角度相交。在实际运用中，你通常也会在那几个小时中移动一定的距离，所以你要通过船速算出你相对第一条位置线所移动的距离，得出一条与第一条位置线平行的新的位置线。真实世界中我们总会出现误差，所以你最好算出三条位置线。如果它们都交于同一点，那不是你实在干得太好，就是实在太走运，不过更常见的情况会是：这三条线围出一个小三角形，水手们叫它"船位误差三角形"（cocked hat）。这个三角形的中点就是你所算出的新位置了。

(9) 这是理想的状态，但我们并不是总能达到这种状态，人性就是如此。一个被现代科学界公认的没能解决的问题就是：错误的结果常常得不到宣传。比如当一个实验得到结果与假设不一致时。费曼在他的"草包族科学"中有力地揭露和批判了学术体系中的这一漏洞，那是他

在1974年加州理工学院的毕业典礼上所做的演讲，也收录在了1985年出版的《别闹了，费曼先生》当中。

(10) by parity of reasoning是以此类推的意思。parody是戏仿、滑稽模仿的意思。——译者注

(11) 这次会议留给我的另一段难忘的记忆就是在威尼斯大运河里泡过的卡尔·波普了。那是在圣乔治岛的船库，波普从汽艇上下来的时候滑了一跤，一脚踩进运河里，在两个身手矫捷的船工拽住他、把他拖上码头之前，河水已经没过了他的膝盖。会议的主办方非常担心，他们急匆匆地赶到酒店想帮90多岁的卡尔爵士换一条干裤子，但是卡尔爵士只带了身上穿的这条裤子。按照日程表的安排，会议在半小时之内就要开始了！这时意大利人的机智体现了出来，在大约5分钟的时间里，我欣赏到了令人难以忘怀的一幕：在一个帕拉弟奥（Palladio）设计的有穹顶和大理石地板的房间的正中央，卡尔·波普爵士在一张小椅子上正襟危坐，至少有6个穿着迷你裙的年轻姑娘围绕着他，她们跪在地上，用吹风机对着他的裤腿不断地吹着。那些吹风机的电线呈放射状一直延伸到墙边，构成了一幅五彩缤纷的人体雏菊画，而我们的卡尔爵士镇定而严肃地坐在花心的位置上。15分钟之后，衣服干了的爵士用拳头锤着演讲台，强调着二元论的观点。

(12) 斯马特（Smart）是聪明的意思，作者把那招称作outsmarting，这个词的本意是“比……更聪明”，是个双关语。——译者注

(13) 《海事法》的复杂是出了名的，里面充满了隐蔽的陷阱和只有海事律师专家才搞得清的例外条款，所以“海事律师业”常用来形容人们利用技术手段逃避责任，或把责任转嫁他人。

(14) 阿克塞尔罗德的囚徒困境锦标赛（Axelrod and Hamilton, 1981; Axelrod, 1984）为“利他性进化”的理论研究领域开创了先河。我在《达尔文的危险观念》（Dennett, 1995, pp.479-480）一书中对囚徒困境锦标赛做过导论性的说明，最近，在世界各地的研究机构当中，各种与之相关的模拟和实验层出不穷。拉波波特简单而精巧地表达了“人不犯我，我不犯人”这一理念，他的工作是迄今所有博弈论研究与博弈模式的源头。

(15) 这里的四条规则是我凭着很久以前拉波波特与我通信的记忆所写的，这些信件现在已经遗失了。最近，塞缪尔·鲁思（Samuel Ruth）向我指出拉波波特法则源自拉波波特的著作《战斗、游戏与辩论》（*Fights, Games and Debates*, 1960）以及他的论文《三种冲突模式》（*Three Modes of Conflict*, 1961），在那篇文章中他清楚地表达了第一条规则，并把这一想法归功于卡尔·罗杰斯（Carl Rogers），还表述了另外几条规则的原型。我的版本可能更加简单、通用一些。

(16) 呆伯特（Dilbert）是美国漫画家斯考特·亚当斯（Scott Adams）系列漫画作品的主人公。——译者注

(17) 不甘示弱的还有哲学家杰里·福多尔（Jerry Fodor），他也把“当然”加上了斜体，并且反复使用，就好像在说：“尝尝这个，怀疑者！尝尝这个，怀疑者！”

(18) 布洛克的英文为Block，心智障碍是mind block，此处是一个双关语。——译者注

(19) 米丽娅姆在网上看到我用了她造的词并和我取得了联系。按她的记忆，当时的场景有些不同，但内在精神还是一致的：“在我们家，这个词有点讽刺意味。它指你为了抬高自己而说的貌似真理的话。”不过，她大方地同意我继续使用她爸爸告诉我的那个版本，并且认可了我对这个绝妙新词所做的重新定义。

(20) 理查德·道金斯的自传《道金斯传》中文简体字版已由湛庐文化策划引进。——编者注

(21) 许多狂热人士会说，动物也能谈论，只是我们还未能发现或转译出它们所用的语言。这是一个具有永久吸引力的观点。但近些年来，有关猿类、鸟类和海豚如何与同类沟通的详尽研究却表明，它们“分享自己想法”的能力是非常有限的。如果它们使用的可以算是语言，那么白蚁就可以成为工程师了，螳螂就有了信仰，狼族可以组织议会。

(22) 杰里·福尔多在他的前沿著作《思想的语言》（*The Language of Thought*）中亮出了这一有力的问题，并引用林登·约翰逊（Lyndon Johnson）挑衅般的说辞“我是你唯一的总统”强调了这个问题所隐含的答案。

(23) 我们常常将这种说法称作精神或意向的整体主义。杰里·福尔多坚决否定这种整体主义，他认为我们完全可以想象出一种只有一种观念的生物（Fodor and Lepore, 1992）。

(24) 其中的重要细节我们将会在第27章中讨论。

(25) 对于那些喜欢想象成为雏菊是一种怎样的体验的人，我表示抱歉。不过我提供了环世界这个概念作为赔礼，它能让你得以公正地对待一朵雏菊真正的能力，而不至于过火。

(26) 哲学家路德维希·维特根斯坦、艺术大师保罗·克莱（Paul Klee）以及批评家维克托·施克洛夫斯基（Viktor Shklovsky）都表达过这个意思。

(27) 让我们列举几个反例来看看它们其中的差别。不论严重与否、形式如何，自闭症最明显的一个症状就是很难形成常识心理。那些患有自闭症的人，比如动物学家坦普·葛兰汀（Temple Grandin），如果想去彻底地了解其他人，只能依赖一些通过自己辛苦努力总结得出的理论。那些我们可以轻松地甚至是不经意间就能看明白的手势、笑容或者状况，葛兰汀必须经过细心观察才能推断出其中的意思。机敏、聪慧的她收集到了足够的线索后，就可以辨别出什么是友好的祝福，什么是威胁恐吓，还能看出人们脸上赞许的神态。她通过自学可以鉴别出承诺和场面话，能讲笑话、能说谎。

《生活大爆炸》中一个众所周知的荧幕形象谢尔顿·库珀（Sheldon Cooper），他实际患有阿斯伯格综合征，这也是一种轻微的自闭症，你几乎能听到他那副科学大脑不停地翻搅，不停地在为当前面对的情景推理出诸多细节，整个过程就像是他在化学实验中分析试管里的催化剂的作用一般。很明显，坦普·葛兰汀和这位虚构的谢尔顿·库珀都有自己的一套心智理论。如果说我们常人也有某种理论，也不会与他们的理论太相像。

(28) 我称之为“意向”立场而不是“理性行动者”立场或者“信念-欲望”立场的原因是，在1971年，哲学家们所做的有关意向性的大部分工作都是包含在“意向习语”的逻辑里的，这种习语记录的是我们对信念、欲望和预期的命题态度，习语中的关涉物常常会引起一些逻辑问题，而且同时也指明了解决方法。一个句子或是一种信念可以表示一些并不存在的东西，比如说圣诞老人；一种欲望也可以针对一个并不存在的东西，比如不老泉或是茫茫大漠中一杯清凉的饮料。意向立场关乎行动者的所思所想，但它不是我们有意去采取的一种立场，它就那么自然而然地、不经意间就发生了。

(29) 我在1983年发表在《行为与脑科学》（*Behavioral and Brain Sciences*）上的专题论文《认知动物行为学中的意向系统：为“过分乐观范式”辩护》（*Intentional Systems in Cognitive Ethology: The ‘Panglossian Paradigm’ Defended*）一文中提供了一些相关细节，其中还夹杂着一连串吸引人的异议和反驳。要想了解有关意向立场的最新阐释以及我对一些常见反对意见的进一步辩驳，请参考拙文《意向系统理论》（*Intentional Systems Theory*, 2009c），本章中多处内容节选自该文。

(30) 这是我在《意识的解释》（*Consciousness Explained,* 1991a）中讨论的中心论题。很明显，有那么一个特殊的地方（我把它称作笛卡尔的剧场）存在的这一观点对很多思想家来说始终具有吸引力，尽管我和其他人已经给出了所有能反驳的理由，但是仍然无法说服他们。

(31) 评论通常不会指出，在那个深受哲学家喜爱的思想实验“缸中之脑”中，活生生的大脑必须有一个身体，并将硬件和多年经验构成的软件相匹配。如果你“恰好”从头开始构建了一个大脑，各元素间随意连接，也就是说，这个崭新的大脑完全是随机连线的，那么有很大的可能，它不会具备任何能力，拥有任何连贯的想法、意图、计划或是明显的记忆，见第35章。也可以参见第32章。

(32) 麦克斯韦妖，是1871年英国物理学家麦克斯韦为了说明违反热力学第二定律的可能性所设想的。是在物理学中，假想的能探测并控制单个分子运动的“类人妖”或功能相同的机制。——编者注

(33) 概念验证是一个有用但又危险的工程学术语：要设计出一个非常精致、非常复杂的小玩意儿，你可以先解决一个“玩具问题”（toy problem），也就是做一个很简单的小东西，让它按照简化的工作原理运行，这就是在“验证概念”。然后需要做的是按照要求让这个小东西的功能充分完善起来，但这只是时间和金钱的问题罢了。这个简单的小东西可以让一个复杂的问题首先在概念上得以“解决”。有时候，“概念验证”就像宣传中那么有效。

(34) 不是每个人都喜欢这种说法。有一本书叫作《神经科学的哲学基础》（*Philosophical Foundations of Neuroscience*, 2003），由神经科学家马克斯·本内特（Max Bennett）和哲学家P.M.S.哈克（P. M. S. Hacker）合著，这本书对我上面的说法持一种鄙夷态度（Dennett et al., 2009）。

(35) 相关的综述和介绍，可以翻阅我的文章《计算机方法的逻辑地理：从东极来看》（*Logical Geography of Computational Approaches: A View from the East Pole,* 1986）。要看决定性的第一手文献，请参考McClelland, Rumelhart, and PDP Research Group（1986）。

(36) 是的，你的枕叶（视觉皮层）还有其他区域的激活模式真的同你眼前所看到的事物形状相似，只是有些扭曲，但另一个问题是，你的大脑里有能看到这些图形的器官吗？也许大脑的某些组织结构有近似可以看到这些图形的功能。无论如何，这需要我们谨慎对待。

(37) 20世纪80年代中期，我和同事乔治·史密斯（George Smith）在塔夫茨大学曾共同教授过一次计算机导论课程，很感谢那个时候他向我引介了寄存器机。乔治·史密斯看到了寄存器机隐藏在教育学里的巨大潜能，并开发出了一种教学框架，我在本章就采纳了这种框架来面向稍有不同的听众，也就是读者您。正是得益于那门共同教授的课程，我后来才和乔治·史密斯在塔夫茨大学建起我们的课程软件工作室。

(38) 习题的答案详见附录。

(39) PC指运行Windows系统的电脑，Mac指运行苹果OSX系统的电脑。——编者注

(40) “可选”意味着没标“可选”的题都是“必选”的！我不是在开玩笑。如果你想从这个思考工具中得到益处，你就必须练习、练习、再练习，直到熟极而流。做那些简单的必选题或许会让你花上一两个小时，但绝对超值。

(41) 埃达·洛夫莱斯，诗人拜伦勋爵的女儿，是一位了不起的数学家，也有很多其他方面的成就。1843年，她翻译了一篇阐释巴贝奇的分析机的意大利语文章，连同她自己的笔记一同出版。她的笔记比她翻译的那篇文章更长，更深入，里面描述了她仔细研究出来的系统，可以用巴贝奇的分析机来计算伯努利数。为此，她经常被誉为第一个计算机程序员。

(42) 认知神经科学和计算神经科学都在研究如何勾画出这些不同的层次。两个领域有各自不同的侧重点，计算神经科学家们更愿意为自己的想法去创造实际的（计算机）工作模型，而认知神经学家则常常倾向于能够描画出有关能力和相互作用的更高层次的模式，他们认为，这些模式对于相对较低的层次也一定适用。类似这样的拉锯战在计算机领域也已经持续了数年。一方面是，人工智能空想家们孜孜不倦地写着他们的实用程序，却没有为任何一项任务程序成功地提供出数据事实。而另一方面，坚强不屈的工程人员们兢兢业业，直到找出任务实际运行的最终代码。除此之外，一切都只算是过眼云烟。他们懒得理那些只忙着推测的空想家们，但这与神经科学领域内不同派别之间的敌视比起来，根本不算什么。记得一位实验室主管（他主要研究神经突中的钙离子通道）曾对我说过，“我们实验室里有这样一种说法，如果研究一个神经细胞，那属于神经科学，而如果研究的是两个，那就属于心理学了。”没有任何赞扬之意！认知神经科学总是吸引着媒体的大部分注意，每个人都关心有关错觉、记忆、意识、讲说和理解方面的新突破，但人们对那成百上千种的神

经调质以及它们的受体，或者是星形胶质细胞、钙离子通道等等却提不起半点兴趣，这让计算神经科学家们嫉恨至极，以致影响到了他们对认知神经科学研究者同僚们的态度。

(43) 我的同事雷·杰肯道夫这样来评论这种语言习得过程悖论：孩子们无意中就能吸收和遵循的一些语法规则，语言学家们费尽九牛二虎之力也未必能表述出来。

(44) 事实上，笛卡尔这一伟大的设想已经触及到了一些非常好的想法。他认为，拉动金属线能为压抑已久的“动物生命之液”（脑髓液）打开一点小小的空隙，让它们奔涌而出，从而产生某种液压动能。真是个不错的力量放大装置！更重要的是，通过这种装置，某些智力活动就完全可以靠机械力实现了：比如无意识的反应。按照笛卡尔的设想，就要踩在一团烈火上时你收住了脚，是因为那种灼热拽动了挂着铃铛的金属线——所有这些都只是大脑的机械反应，不需要理智！

(45) 说点有关专家系统的题外话：CYC项目（Lenat and Guha, 1990）无疑是最令人印象深刻的人工智能系统，它是思想语言这类想法的执行系统，一部百科全书（enCYClopedic），它拥有人工编制的海量数据库，并配有专属推理引擎负责管理该数据库。目前，CYC项目的发展已经超过四分之一个世纪，参与人数众多，但就其设计而言，它力所能及的范围依然是冷冰冰的非生物、非心理的领域。参见Wikipedia上关于CYC的精彩介绍。几乎可以肯定，迈克有关红发人的看法不是一个用脑语表达的公理化的红发人微型理论，也未配备类似于CYC的那种大型数据库。我们尚不了解通过这类东西能做些什么。原因是，除了一些非常简易的模型，例如著名机器人制造专家罗德尼·布鲁克斯（Rodney Brooks）及其同事的类昆虫式包容架构（Brooks, 1987），我们目前仍未对此有过直接研究。布鲁克斯的Cog项目研制了一种类人机器人，这个项目的一个主要理论贡献在于：将这些高度非命题式的内容结构模型套用在人类心理上（Dennett, 1994b）。

(46) 在一开始创造这个直觉泵时，我或许不该使用这种过时的俚语。但我是不得已而为之，因为这种东西曾是一种流通量颇大的货币，所以让我继续用它吧。用“two bits”表示25美分有着模糊不清的起源，这让我们回想起旧时的西班牙银币八里亚尔、达布隆或其他海盗时代的遗物。

(47) 像你祖母常常用来熨衣服的那种尖头熨斗，就可以是一个相当不错的门挡。如果你想在旧货店或者“收藏品”网站淘一个的话，一定要确保它不是赝品。市场上有些不过是铁制的门挡，只是铸成了旧式尖头熨斗的样子。未来一百年也许就有人会造出稀奇古怪的门挡，看上去就像双币机，在硬币还能当钱用的年代，你的祖父还常常用它。你懂的。

(48) 哲学家希拉里·帕特南（Hilary Putnam, 1975）多年前发明了孪生地球这一思想实验，我的直觉泵细致地模仿了相关细节。我的双币机的故事只是煞费苦心地重新设置了帕特南直觉泵上的旋钮。在过去35年间，哲学家们讨论过几十种甚至可能是上百种各式各样的变体。在最初的叙述中，帕特南选取了地球上的水和孪生地球上的水（其组成为XYZ，而非H_2O）作为例子，但这会引入一些与我们的直觉泵不相关的复杂情形，所以，在这段叙述中，我们选用了犸——就好比双币机中的巴波亚硬币。

(49) 特库姆塞·菲奇（Tecumseh Fitch）在其重磅论文《纳米意向性：为内在意向性一辩》（*Nano-Intentionality: A Defense of Intrinsic Intentionality*, 2008）中提出，最早进化出“内在”意向性的实体是真核细胞，而非其祖先原核细胞。按他的说法，这是因为它们自我保全的天赋远较其祖先高超。他对单个细胞的自治性（能动性）的重视强烈影响了对我的小人儿式机械主义的各种修正（见第20章）。不过，对于他止步真核细胞的打算，我不表苟同。原核细胞与真核细胞一样，不能“被机器替代”；能动性可以一路下落直至蛋白质，最后就是自私的基因。

(50) 对本段坚持不懈而又激情四溢的反对来自福多尔和皮亚泰利-帕尔马里尼（Piatelli-Palmarini）的《达尔文弄错了什么》（*What Darwin Got Wrong*, 2010）。本书和那本书之中，肯定有一本是在大放厥词。至于为本书观点所做的详细辩护，见本书第五部分及拙作

《理性的进化》（*Evolution of Reasons*）。

(51) 自1984年以来，露丝·米利肯发表了一系列才华横溢的著作，她较我更为细致地发展了这一主张。欲了解最新动态请参阅《米利肯和她的批评者》（*Millikan and Her Critics, Ryder et al.*, 2013）。

(52) 哲学家约翰·豪格兰德（John Haugeland, 1981）说，人工智能的第一原则是，“你若搞定句法，语义将自己搞定自己”。人们对这一口号有着不同的理解。最初的版本抱着过高的希望，鼓励人们去检索大数据库：这种数据库是一个世界知识的公理化形式体系，可以由某个纯粹句法的推理引擎维持并加以利用，CYC就是最好的例子。这在大多数专家眼中已被视为是不可行的，但我们不妨将该口号看作是对以下想法的优秀表述：大脑是一种计算机器，并因此是一部句法引擎，得益于其设计，它大致在做语义引擎的工作。

(53) cow shark是六鳃灰鲨（Hexanchidae）的俗称，作者的例子可能源于此。——译者注

(54) IBM超级电脑沃森的及时出现几乎将我的科幻故事变成了科学事实。如果乐意，你可以想象把沃森放入A盒子，而B盒子包含一个瑞典版的沃森，由Bo独立开发。当我首次发表该思想实验时（Dennett, 1995a），我所能说的不过是：“想知道真实世界中的一个例子，看看道格拉斯·莱纳特（Douglas Lenat）在微电子与计算机技术公司（MCC）的巨型CYC项目（Lenat & Guha, 1990）。”参见第28章关于CYC的脚注。沃森所采用的人工智能方案早在1995年还是难以想象的一件事，但它现已取得了巨大的进步。不同于CYC主要由手工编码，沃森可以在无人值守的情况下从互联网上获取事实信息，并能使互联网上那些可用数据的统计特性发挥强大作用。沃森和CYC均能以各自不同的方式近似理解它们数据库中的数据，它们对数据的理解远比其他拥有大型数据库的计算机强得多。

(55) 在我一开始写这个直觉泵的时候，ASCII码（美国信息交换标准码）几乎是所有文字处理、电子邮件以及互联网语言的标准格式。现在，它已经被一种扩展了的逆向兼容格式UTF-8所替代，UTF-8意为8位的通用字符集转换格式。也就是说，ASCII码是UTF-8编码格式的一部分。

(56) 鲁布·戈德堡机械（Rube Goldberg machine）是一种设计得过度复杂的机械组合，以迂回曲折的方法去完成一些其实非常简单的工作，例如倒一杯茶或打一只蛋等。——译者注

(57) 本注只写给哲学家：已有人论证，我关于“实模式”（Real Pattern, 1911b）的解释是内容的副现象论（epiphenomenalism）。这就是我的回复。

(58) 一旦他们猜到了这种假设，火星人就能参与到蒯因的彻底翻译任务中，但是他们的任务将变得愈发困难，因为他们不能对“信息来源”，也就是A盒子或B盒子发问，比如拿起一个物体问，这是“球”？“铅笔”？诸如此类。

(59) 英文字母N和O对应的摩尔斯码分别是-·，---。——译者注

(60) 在解释进化史的历史事实时，我们可以得到同样的教训：即便你能细致入微地描述每只长颈鹿生活过程中的每一个因果事实，也永远无法解释那些显而易见的规律，比如，长颈鹿为什么长着长长的脖子，除非你上升一到两个层面追寻大自然母亲的理由，追问“为什么”。（更多讨论见本书第五部分。）

(61) 我有预感，一个恰当的有关意义的科学理论将呼之欲出，它会取代有效的老式人工智能的布尔式刚性逻辑结构，代之以一个更灵活的、基于模式搜索算法的贝叶斯统计或概率网络，但是要实现这些想法尚需时日。

(62) 西班牙乌蝇（Spanish fly），又称斑蝥，据说是臭名昭著的萨德侯爵发现了它的催情功能；传闻硝石（主要化学成分为硝酸钾）也有类似功能。——译者注

(63) 一些评论人士认为，我将进化论同万能酸两相对照的灵感来自DNA，因为DNA是脱氧核糖核酸的缩写，但我想表达的内容更为普遍：在我们的星球上，DNA并不是进化的唯一媒

介，而且谁知道宇宙间是否还有其他进化图景呢？

(64) 克雷格·文特尔的著作《生命的未来》中文简体字版已由湛庐文化策划引进。——编者注

(65) 博尔赫斯选了一组稍微不同的数字：每本书有410页，每页有40行，每行有80个字符。所以他的每本书的总字符数为1 312 000，和我的设置1 000 000个字符大致相当。只是为了便于操作，我才选择了一个整齐的数目。博尔赫斯所采用的字符集只有25个元素，这对于大写的西班牙语来说已经足够了，而且他的标点符号只有空格、逗号和句号，但这并不适用于英语。所以我选择了一个更为宽裕的字符集，可以容纳100个字符，使它足以包含所有使用罗马字母的语言中的大小写字母以及标点符号。

(66) 斯蒂芬·霍金（Stephen Hawking, 1988）故意这样说："在我们能观察到的宇宙区域中，有一万万万万万万万万万万万万万万万万万万万万（1后边80个零）个粒子。"对于可观察的宇宙，迈克尔·丹顿（Michael Denton, 1985）估计存在着10^{70}个原子。曼弗雷德·艾根（Manfred Eigen, 1992）算出宇宙的体积为10^{84}立方厘米。

(67) 巴别图书馆是有限的，但奇怪的是，它包含了所有合乎语法的英语句子。但这是一个无限集，而图书馆是有限的！尽管如此，任何一个任意长度的英语句子都可被拆分为500页一个的字符块，散布在图书馆的某个地方！这怎么可能？——有些书可能会不止一次地被使用。最极端的情况是最容易理解的：因为某些书仅仅包含单个字符，而有些书纯粹是空白的，重复利用这一百本书就能创建出任意长度的任何文本。正如蒯因在其颇具信息量且不失风趣的文章《万有图书馆》（*Universal Library*, 1987）中指出的那样，如果你利用这一策略，即反复使用这些书，并将所有内容都翻译成可供文字处理器使用的ASCII码，那么你就可以将整个巴别图书馆存储在两本极薄的书中，其中的一本印着一个"0"，而另一本印着一个"1"！蒯因还指出，心理学家西奥多·费希纳（Theodor Fechner）早在博尔赫斯之前就提出了万有图书馆这一奇思妙想。

(68) 蒯因（Quine, 1987）出于同样的目的发明了"超宇宙的"（hyperastronomic）一词。

(69) 这部小说名叫《笑声之谷》（*The Vale of Laughter*, *De Vries*, 1953）。文中继续写道："请随意。无论白天黑夜，随时给我打电话都可以。"德弗里斯相当于发明了一种游戏，看看一个印刷错误能带来多大的效果，无论这种效果是否对文意有害。

(70) 《白鲸》的原文是"Call me Ishmael"（叫我以实玛利），与该句相差一个逗号。——译者注

(71) 古希腊和罗马戏剧中用以完结剧情的神，其字面意思为"机关里跑出来的神"。在演出时，扮演神的演员会被某种装置送入舞台，故名。——译者注

(72) 好吧，其实也不是完全不可能。轨道运行与地球自转同步的地球同步卫星就是一种真正的、非奇迹的天钩。它们值得大笔投资的价值在于，我们常常很想让某些东西，如天线、摄像仪以及望远镜悬停在高空中。不过，将卫星用作提升重物是不切实际的，因为它们必须安置在极高的高空。人们曾经认真地探讨过这种想法。然而事实证明，提升所用的绳索即使由最强人造纤维制成，其顶部的直径也必须超过100米，越往下绳索可以越细，直至变成几乎看不见的鱼线。这还只是为了承受其自身的重量，更谈不上任何有效负载。即便你能织出这样的绳索，你也不会希望它从轨道上掉落到下面的城市中去！

(73) 我很高兴地看到，2012年8月6日，好奇号火星探测器在"天空起重机"的帮助下在火星着陆。尤其令人高兴的是，这个着陆设备被称作天空起重机而不是天钩，因为它是一件工程杰作而非一个奇迹。

(74) 2006年8月24日，在布拉格举行的第26届国际天文学联合会通过第5号决议，将冥王星划为矮行星，自此冥王星被从太阳系大行星行列中除名。所以，目前太阳系只有八颗大行星。——译者注

(75) 美国“西大荒”年代赫赫有名的“牛仔之城”，位于堪萨斯州，美国西部。——译者注

(76) 注意，线粒体夏娃和Y染色体亚当的遗产有一处重要差异：无论男性还是女性，我们所有人的细胞中都有线粒体，它们均来自我们的母亲；如果你是男性，那你就拥有一条Y染色体，它来自你的父亲，而几乎所有的（但也不是全部）女性根本没有Y染色体。

(77) 弗林是位出生于澳大利亚的演员，活跃于20世纪30、40年代的好莱坞。他塑造了诸多外表英俊、风流倜傥的叛逆者形象，为了正义和美女的芳心不停的战斗，但在私生活方面饱受诟病。——译者注

(78) 不考虑逻辑上的可能性，来看看真实的历史，我们注意到，历史上有一些男人试图在其种群中垄断生殖，这种垄断有时到了令人瞠目结舌的地步：国王或强人会霸占上百名女性并使之受孕。这些人很可能对我们人类的基因池做出了很大的贡献，无论是好的贡献还是不好的贡献。

(79) 美国作家戴尔·彼得森（Dale Peterson）让我注意到这样一个事实：黑猩猩的石锤经过一段很长时间的持续使用才得以磨光。这就为一系列事件补上了一个恰当的“缺失环节”：一开始，你漫不经心地捡起一块石头来砸坚果，接着，你在目所能及的范围内选择具有最佳形状的石头来完成这项任务，接着你又在更广大的区域内寻找一块更合适的石头，然后，你通过肉眼和视觉记忆注意到，你中意的石块变得越来越顺手，最后，当你为预想的任务打造石器时，你已经变得非常老练了。

(80) 对于非生物循环影响生殖细胞出现的可能性的作用，我将在即将发表的新文章《理性的进化》里探讨有关细节。

(81) 这段德文是：Freunde, Römer, Mitbürger, gebt mir Gehör! Ich komme, Cäsars Leiche zu bestatten, nicht, ihn zu loben.这是莎士比亚的戏剧《裘力斯·凯撒》中，安东尼在凯撒葬礼上著名演讲的开场白，中文翻译如下：“各位朋友，各位罗马人，各位同胞，请你们听我说：我是来埋葬凯撒，不是来赞美他。”——译者注

(82) 一位电脑黑客生造了“thinkos”一词，用来指称语义层面而不是句法层面上的编码错误。忘记半个括号是排印错误，而忘记声明某个局部变量则是思维错误。在对人类活动做出的语义解释或意向解释，以及在明确的正误或简繁标准中，均可能出现思维错误。形容一位淑女庸俗（meretricious）而非值得称赞（meritorious）就是思维错误，而不是排印错误（Dennett, 2006b）。

(83) 指棋子中的“车”和“后”。——译者注

(84) 粗略地说，MIDI之于音乐好比ASCII码之于计算机书写系统，是计算机程序与外部世界之间的通用语。

(85) 鲁本斯是巴洛克美术的代表人物。他修筑了极为奢华瑰丽的画室大夏，在这里，成群的学徒、合作者，以及各类艺术家簇拥在他的周围，协同创作，当然，这些作品都冠以“鲁本斯”之名。到1620年前后，其画室有超过100名学徒，俨然一个高产的艺术工厂。因此，甄别其作品真伪的争辨至今不绝。——译者注

(86) 一位读者在看完这部分草稿后觉得我是在瞎编。如果你也有这样的怀疑，请在庄重的“斯坦福哲学百科全书”（*Stanford Encyclopedia of Philosophy*）上查阅《哲学僵尸》这篇冗长又严肃性十足的文章吧。

(87) 这又是一处某些哲学家会指责我的地方，他们认为我混淆了认识论问题与形而上学问题：“不要问我们是怎么知道一个人是僵尸的！要问就问，僵尸是怎么回事！”离开了认识论的形而上学顶多是幻想。

(88) 候世达从塞尔文章中的一个用语“几张纸”入手，向人们细述了塞尔如何将他那个软件的规模和复杂程度低估了数个量级。在我们这本书中，候世达对塞尔的评论稍带批判色彩，这

招致了塞尔1982年在《纽约书评》中的一顿恶骂，原因是候世达手一滑，在自己的评论中将塞尔文中的“几张”写成了“几片”，尽管我们确实是如实转载的文章，但塞尔却硬说我们误传了他的论述！如果他说的情况属实，如果一个无意的微小变动真的可以改变整个机制的运作，那在某种程度上不正好证实了我们的观点。如果一个那么小的细节就能从根本上左右一个思想实验，那么人们更应该多留心这类实验“泵”出的直觉了。

(89) 侯世达在《心我论》中很专业地表述道：中文屋这个直觉泵至少有五个旋钮，他会依次转动这些旋钮，从而派生出一系列不同的思想实验，引出不同的结论。在结论中，我写到：

> 直觉泵上每一个旋钮的变动都会让相关叙述变得稍有不同，这时，原有的问题可能会暂时隐藏，而不同的寓意则会浮出水面。所以，我们要仔细审查，确定哪个或哪些版本的叙述是可信的。如果直觉泵的创造只是为了尽量简化而不是为了压制不相关的干扰，那么我们最好不要轻信由它引出的那些结论。（Hofstadter & Dennett, 1981, p. 460）

(90) 大家都认为这与笛卡尔的说法类似。1637年，笛卡尔在《谈谈方法》中写道，将一台机器和一个装有无形灵魂的人区别开来的最好方法是与他们对话。

> 这种机器的确可能存在：它可以说话，会恰当地说出自己的行为状态或者引起自身机体变化的那些物体。看起来，好像你碰它一下它它就会问你想说什么，一受伤它就会哭泣似的，诸如此类。但面对同样的问题它只能提供同一个答案，就像那些最愚蠢的人一样。

而图灵直觉泵的发明是否受到过笛卡尔这个直觉泵的启示，现在我们已无从考证。

(91) 这些问题有：你怎么知道谁是主人，谁是替代者？当你发现自己近期的记忆中充斥着几小时到几天的空白时，你难道不惊恐吗？

(92) 在艾略特的一部诗歌剧《大教堂凶杀案》（*Murder in the Cathedral*）中，当问及过去发生在贝克特身上的事件时，他回答：

你说起的季节已经过去。

我只记得不值得遗忘的事。

(93) 作为意识科学研究的选择方法，我为他者现象学的辨护引起了相当大的争议。一些研究人员认为这是对使用科学研究意识的条件的一种改变游戏规则的澄清，有人认为这只是在重复显而易见的东西，还有一些人则继续反对。想要了解更好的理论，请查阅书后的本章来源。

(94) 保罗·布卢姆，著名认知心理学家。他的著作《快感：为什么它让我们欲罢不能》《善恶之源》中文简体字版已由湛庐文化策划引进。——编者注

(95) 如果你有所怀疑，请参见Vohs和Schooler于2008年的一项研究及后续相关实验证据。

(96) 我在1987年出版的《意向立场》（*The Intentional Stance*）的第9章中更加细致地阐述了这种时空变换的理论含义。

(97) A.K.杜德尼（A. K. Dewdney）在《平面宇宙》（*Planiverse*）一书中以一个完全不同的角度研究了二维世界的物理学和工程学，他对埃德温·A·艾勃特（Edwin A. Abbott）在小说《平面国》（*Flatland*）中提出的设想做出了巨大的改进，而他的工作本身也是一个极为了不起的思考工具。

(98) 影片《公主新娘》（*The Princess Bride*）当中有一个精巧的段落描绘了一场失控而复杂的互相揣度。在剧中，维奇尼（华莱士·肖恩扮演）想让维斯特雷（加利·艾尔维斯扮演）猜“哪个杯子里有毒药”，结果反而害了自己。

(99) 据英国王室官方网站称：“今天，英国王室仍然保留着对所有开放水域内的野生天鹅的所有权，但女王只在泰晤士河及周围支流行使所有权。该所有权与名酒酒商公会和染织商名

家公会共享，早在15世纪这两家公司就获得了王室授予的这项权利。当然，如今的天鹅们已经不会再被吃掉了。”

(100)　《自由意志》的中文简体字版已由湛庐文化策划引进。——编者注

(101)　欧洲中世纪神话中的魔鬼。它最著名的形象是歌德在小说《浮士德》中所塑造的一位诱惑浮士德的魔鬼。——译者注

(102)　哈耶斯的工作启发了他人。一本相关的文集中的一篇文章指出：“朴素物理学本身就是一个模糊的术语。这只是糟糕的物理学吗？是心理学、人工智能、或者物理吗？”（Bobrow, 1985, p.13）我认为朴素物理学试图使常识映像的物理部分足够严格以支持自动演绎推理。

(103)　关于这本书的第30章中出现的谜题，答案1、5、6、7，分别是slop、love、even、Dern或者smut、hope、iron、Penn。其中两名影星分别是Laura Dern和Sean Penn。

(104)　国际象棋是chess，丹尼特把chess中间加了一个m，变成了chmess，任何语言里都没这个词语，我把它翻译成了“像棋”。——译者注

CHEERS
湛庐

为什么需要生物学思维

OVERCOMPLICATED

洞悉复杂世界的思考方式

TECHNOLOGY AT THE LIMITS OF COMPREHENSION

SAMUEL ARBESMAN
[美] 塞缪尔·阿贝斯曼 著
贾拥民 译


圣塔菲书系 Santa Fe Institute Series


四川人民出版社

版权信息

本书纸版由四川人民出版社于2019年5月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：为什么需要生物学思维

著者：塞缪尔·阿贝斯曼

电子书定价：62.99元

Overcomplicated: Technology at the Limits of Comprehension by Samuel

Arbesman

Copyright © 2016 by Samuel Arbesman. All rights reserved.

此书送给

将在一个奇异世界中长大的

阿比高尔和内森

请永远不要失去对这个世界的激情

引言　认识复杂系统

2015年7月8日，我还在撰写本书的时候，美国联合航空公司的计算机系统出了故障，所有飞机被迫停飞。同一天，纽约证券交易所的计算机系统也遇到了问题，所有交易被迫中止。《华尔街日报》的网站也崩溃了。接二连三的系统问题令无数人惊慌失措，但是没有人知道到底发生了什么。Twitter上谣言四起，许多人都在猜测：美国是不是遭到了网络攻击？

但是，这些似乎并不是有预谋的网络攻击。“罪魁祸首”更有可能是问题很多且尚未被人完全掌握的软件系统。[1]一位安全专家在评论该日事件时所说的话可谓一针见血：“这些都是极其复杂的系统，出错的方式多种多样，其中有很多是我们目前完全无法理解的。”[2]事实上，对于这些复杂得令人难以置信的技术，说我们根本无法尽知其出错方式，其实已经是在有意地轻描淡写了。

人类的技术，从网站到交易系统，从城市基础设施到科学模型，甚至是为大型企业提供配套服务的供应链系统和物流系统，都已变得过于复杂且相互交错。在很多情况下，就连当初构建、每天使用和维护它们的人也无法完全理解了。

在《创造力差距》（*The Ingenuity Gap*）一书中，加拿大政治学家、生态学家托马斯·霍默-狄克逊（Thomas Homer-Dixon）描述了他于1977年在法国斯特拉斯堡（Strasbourg）参观高能粒子加速器时的情景。[3]当时，他问一位在那里工作的科学家，想知道是否有人明白整台加速器到底有多么复杂，结果得到的答案是：“没有任何人能够完全搞懂这台机器。”后来，霍默-狄克逊回忆说，这个答案令他深感不安。是的，我们理应感到不安。从那时起，粒子加速器以及人类所构建的其他一切事物的复杂程度，几乎都在不断增加。

技术的复杂性一直在增加。以铁路为例，铁路出现后，轨道网络和调度系统应运而生，以保证列车能够安全且准点运行。因为要对遍布整个大陆的无数列车进行调度，所以美国的标准化时区制也得到了发展。在铁路技术诞生之前，或者说在其复杂情况接踵而至之前，标准化时区制本不是必需品。

然而时至今日，技术复杂性已日益接近一个临界点。计算机的出现

又给环境平添了许多激进新颖性（radical novelty）[4]，这是计算机科学家艾兹格·迪科斯彻（Edsger　Dijkstra）提出的一个术语。计算机的硬件和软件比之前的任何事物都要复杂得多：一个程序可能包含数百万行计算机代码，微芯片内集成微型元件的尺寸也已经小到接近原子量级。现在，计算机已经被应用到了汽车、手机，以及金融市场等领域，而不断增加的技术复杂性也已经到了令人费解的程度。

近年来，科学家们甚至还逐渐认识到，技术已与自然密不可分。研究地球岩层的地质学家已经在正式考虑这样一个问题：是否有足够的证据支撑我们把当下这个时代命名为“人类世”（Anthropocene）[5]，也就是人类的时代。无论正式名称是什么，我们所制造的人工系统与自然世界之间的亲密关系都意味着，人类的每一个行动都比以往任何时候更有可能引发意想不到的后果，其影响不仅将涉及生活的方方面面，还将波及地球的每个角落，甚至在某些时候会超出地球本身。

从整体上看，技术和基础设施已经变成了极其复杂的、物理化和数字化的系统，好似血液循环系统那般，牵引着地球上所有原材料的流动，同时“排放”出道路、摩天大楼、芸芸众生，以及化学废水。技术加速了地球的“新陈代谢”，而这个过程是在一场极其复杂的“材料之舞”中完成的，这一切甚至改变了地球表面的光芒。

面对这种复杂性，我们往往会情不自禁地发出心难两用的感慨。一方面，我们构建出了这些令人难以置信的复杂系统，这无疑是值得自豪的，尽管它们也许永远无法如预期那般正常运行，但是它们作为无比复杂的大型系统，本身就是值得赞叹的。另一方面，我们在技术领域所付出的所有努力，几乎都在使我们远离优雅和简单，并且让我们无法抗拒地一步步走向某个极度复杂、完全无法预料的世界。

这段“旅途”的终点，其实已隐约可见，那是一个几乎可以自成一体的技术生态系统，一个超越了人类认知和理解能力的极限世界。正如2013年9月《科学报告》（*Scientific　Reports*）所刊发的一篇论文指出的那样，现在，一个完整的“人类不可能及时应对的新机器生态”已经出现了。[6]需要注意的是，这篇论文所讨论的范畴还仅仅是金融领域而已。在股票市场上，各种机器通过各种形式进行着交互，从根本上说，这与利用不同算法进行交易没什么两样，而人类则成了局外人。[7]

本书的观点是，如今存在着这样一种趋势和力量：无论人们做什么，都会使技术日趋复杂且变得难以理解。这种趋势和力量意味着，2015年7月8日那天所发生的事件将会越来越多。人们原本认为非常可靠

的系统会发生无法解释的故障，甚至崩溃！

作为一名研究复杂性的科学家，我花了很多时间研究生活中日益增加的复杂性。我早就注意到，在巨大的复杂性面前，普通人一般会做出两种反应，这两种反应都因各自的极端化而有失偏颇：第一种反应是对技术所导致的复杂世界满怀恐惧；第二种反应是对技术过度崇敬且深信不疑。

恐惧是一种非常自然的反应。长久以来，人类一直被“淹没”在各种各样的末日预言中：大型机器会变成杀手，超级智能计算机即将出现，程序无法保证自动驾驶汽车不会撞死行人，等等。这些技术都非常复杂，即使是专家也无法完全理解，碰巧它们又相当可怕。这样一来，我们在接近技术、使用技术的时候便会警觉不安、惊恐不宁。

就算对技术系统本身并不畏惧，许多人还是会对已经形成包围之势的算法和技术如履薄冰，甚至感到厌恶，不愿直面技术所拥有的惊人力量。当亚马逊或奈飞推送购买建议时，我们无从拒绝；当敲下的文字被输入法自动纠错时，我们恼怒不已，这些都是我们的切身体会。甚至，在面对应用程序所推荐的从一个地方到另一个地方的最佳路线时，许多人偏要反其道而行之。毫无疑问，这种“算法厌恶”现象反映出了许多人共有的一种情绪，而这种情绪其实就是“技术恐惧”[8]的表现之一，只是不那么严重而已。

与此同时，还有些人走向了另一个极端：对技术过度崇敬。在面对某个行为奇异的复杂事物时，很多人最终都会选择到宗教中寻求慰藉。当领教了谷歌公司“大脑”的“博学”[9]，也就是对用户需求的精准预测之后，抚摸着心爱的最新款苹果手机时或访问一个庞大无比的数据中心时，我们的内心可能会蠢蠢欲动，感觉自己像是走进了一座庄严雄伟的大教堂。这时，崇敬之心油然而生。

然而，这两种反应，无论是来自专家的，还是来自非专业人士的，皆不恰当，或者说，皆是非生产性的。在并不值得我们深信不疑的系统面前，无论是恐惧也好，崇拜和敬畏也罢，都会阻碍我们选择恰当的面对方式。若无法采取切实到位的措施，我们就有可能失去控制权，而这会带来意想不到的、甚至是毁灭性的后果。当系统再度崩溃时，这种“理解不足”所导致的问题可能就不会再像《华尔街日报》事件中那样，仅仅只是读者无法在网站上阅读文章那么微不足道了。故障可能会发生在电网、银行系统，甚至医疗领域中，而且不会自行消失。我们现在如果不正视它们，未来定会陷入险境。

技术无处不在，既不古老也不深奥，说到底，它源自某些完美的无限遐想。[10]纵然技术是十分混乱且不完美的，但它却很“平易近人”。我们所需要的是应对各种情况的策略。

我写这本书是为了帮助人们在上述两个极端反应，也就是恐惧和崇敬之间找到一条正路，并为接近技术、理解技术和掌控技术探索出方向。那是一个乐观的方向，能够改变对这些系统的思考方式，让我们不会陷入恐惧或崇敬，不会陷入“六神无主”。

想要朝这个方向出发，就得向技术“妥协”，而这又要求我们得先培养出一种面对技术系统时的舒适感，尽管我们从未完全理解它们。这是一种谦卑的舒适感，既要容忍不确定和不完美，又要不断努力探索，以期逐渐了解更多。正如即将看到的那样，除了其他方面，我们在探索这个方向时，还需要借鉴科学家研究生物复杂性时所采用的方法。

尽管人类过于依赖技术系统，而这些系统又过于复杂，但我仍然相信，人类终有希望掌控自己所构建的系统。

本书旨在说明为什么系统会过于复杂，以及我们应该怎样理解和掌控过于复杂的系统。



扫码下载“湛庐阅读”App，

搜索“阿贝斯曼”，

获取作者演讲视频及更多精彩内容。

目 录

引言 认识复杂系统

01 欢迎来到这个纠缠的时代

连接和反馈是复杂系统的典型标志。技术系统已经变得如此复杂，以致每个专家都仅知道其中的一个部分，没有人能够完全理解整体。我们已经从“启蒙时代”迈进了“纠缠时代”。

什么是复杂系统
从“启蒙时代”到“纠缠时代”
抽象的局限

02 复杂系统形成的4个原因

技术系统变得越来越复杂的主要原因是“吸积”和“交互”。随着时间的推移，系统中不断加入更多的组成部分，部分之间也增加了越来越多的连接。“必须处理的例外情况”和“普遍的稀有事物”也让技术系统变得愈加复杂。

原因1：吸积
原因2：交互
原因3：必须处理的例外情况
原因4：普遍的稀有事物
越来越多的复杂系统

03 为什么复杂系统越来越难以理解了

对于大多数人来说，记住一个7 位数就相当不容易了。当复杂系统的组成部分和连接数量猛增时，即便是专家也会望而生畏。我们个人的知识储备，与理解复杂系统所需要的知识相比，存在着根本性的冲突。

力不从心的大脑
认知的极限
最后一个无所不知的人

04 令人费解的bug

bug并不都是能够找到确切起因的那些错误，在以连接和交互为特征的复杂系统中，经常会出现一些令人费解的bug。尽管我们对这些bug 没什么好感，但它们却是这个纠缠时代中无法回避的存在。

并不是所有bug都能被消除
从错误中学习

像生物学家一样思考

05 为什么需要生物学思维

复杂的技术系统更接近生物学系统，因此，用生物学思维思考复杂技术是个不错的选择。为了从整体上理解系统，我们也会忽略掉一些细节，这时，物理学思维才是首选。我们真正需要的是经过物理学思维锤炼的生物学思维。

复杂的技术系统需要生物学思维的3个原因
技术领域的“生物学家”
当物理学遇见生物学
复杂性科学的视角
思维方式的进化
我们需要通才

06 生物学思维是理解复杂世界的一把金钥匙

认识复杂系统的正确态度是：对于难以理解的事物，要努力克服我们的无知；一旦理解了某个事物，也不会认为它是理所当然的。谦卑之心，加上迭代的生物学思维，就是洞悉复杂世界的正确方式。

不要被表象迷惑
以欣慰感看待不理解的事物
谦卑之心＋迭代的生物学思维

注释
延伸阅读书目
致谢
译者后记

OVERCOMPLICATED

Technology at the limits of comprehension

连接和反馈是复杂系统的典型标志。技术系统已经变得如此复杂，以致每个专家都仅知道其中的一个部分，没有人能够完全理解整体。我们已经从“启蒙时代”迈进了“纠缠时代”。

01 欢迎来到这个纠缠的时代

连接和反馈是复杂系统的典型标志。技术系统已经变得如此复杂，以致每个专家都仅知道其中的一个部分，没有人能够完全理解整体。我们已经从“启蒙时代”迈进了“纠缠时代”。

1986年初的一个冬日，距离“挑战者号”航天飞机爆炸事件发生还不到一个月，著名物理学家理查德·费曼（Richard Feynman）在调查委员会的听证会上公开了一份调查报告。[1]费曼非常冷静地揭示了导致“挑战者号”航天飞机起飞后不久就爆炸解体的原因。问题出在用于密封固体火箭助推器各部分之间缝隙的O型圈上。O型圈其实是一小片橡胶，放入冰水后会失去弹性，也就是说，这种小部件对温度变化非常敏感，所以无法保证密封的牢固性。这些O型圈很可能就是这起导致7名机组人员丧生灾难的罪魁祸首。

与此事件如出一辙的是发生在汽车业中的另一事件。2007年的一天，琼·布考特（Jean Bookout）驾驶着一辆生产于2005年的丰田凯美瑞，行驶在俄克拉何马州的一个小镇上。突然，她的车开始加速，并失去了控制。[2]她踩下了刹车，但没有任何作用；她拉起了紧急制动器，轮胎在道路上留下了触目惊心的摩擦痕迹。然而，这些努力都无济于事，汽车最终狠狠地撞到了路堤上。布考特受了重伤，车上的乘客、她的朋友芭芭拉·施瓦茨（Barbara Schwarz）则不幸遇难。

并不是只有布考特一个人遇到这种怪事。事实上，多年以来，在丰田公司生产的汽车中，有多种型号的汽车都曾出现过此类奇怪且危险的故障：汽车在行驶过程中会“违背”驾驶员的意愿保持匀速，甚至还会“擅自”加速。这种意外加速事故已经导致多人伤亡。业内专家分析出了这类事故的几个导火索：驾驶员操作失误，垫毡顶住了油门加速踏板，油门加速踏板黏滞未能及时复原，等等。但是，这些原因并不能解释所有的意外加速事故。在受事故影响的车型中，因为垫毡或油门加速踏板不合格而被召回的并不多，而且我们也没有理由认为，丰田汽车的驾驶员比其他汽车的驾驶员更容易犯错。

无奈之下，丰田公司请来了嵌入式软件专家迈克尔·巴尔（Michael Barr）追查原因，并向他开放了专有数据库以及向来严格保密的软件代码。在6位经验丰富的工程师的协助下，巴尔几经周折终于找到了问题的根源。与此同时，计算机科学家菲利普·库普曼（Philip Koopman）也通过对公开程序的仔细分析，找到了事故发生的大致原因。[3]两位专家都认为，对于意外加速事故的发生，丰田发动机软件系统的过度庞大和极度复杂，以及糟糕的设计都应该承担相应的责任。我们无法将事故责任明确地归咎于某个设计或部件出了错，毕竟，这里面存在的问题盘根错节，而且这些问题还会引发汽车软件系统与外部机电系统之间的巴洛克式结构（baroque structure）的大规模交互。无论是单独看来，还是整体看来，这个系统的极度复杂性让我们很难理解这些相互作用、相互影响的部件的深层问题和缺陷。有依据表明，丰田公司本来完全没有必要搞出一个如此复杂的系统，至少在现有情况下，不必如此复杂，它们显然不够小心。[4]

这个“诊断结果”与我们熟悉的“技术失败”并无二致。费曼找到了导致“挑战者号”航天飞机爆炸的单一原因，有所不同的是，我们无法确定导致丰田汽车出现事故的单一原因。无论如何，如果部件堆积得越来越多，设计也越来越复杂，那么灾难便只会越来越多。

事实上，即使找到了导致失败的单一原因，但在极度复杂的系统中，那很可能也只是掩人耳目的东西罢了。1996年，阿丽亚娜5型火箭在发射升空39秒后爆炸，火箭上搭载的4颗卫星均毁于一旦。[5]事故分析表明，发生爆炸的原因是，火箭在新环境下使用了一些较为陈旧的软件代码。但是，根据霍默 - 狄克逊教授的说法，没有任何一个承包商被追责。这次爆炸并不是某个决策失误所致，而是整个发射系统的极度复杂性所致。其他类似的灾难还有三哩岛核电站事故，[6]尽管有关方面也

找到了一个原因，但是究其根本，仍然得归咎于系统的庞大和极度复杂，而非某个部件或决策出了错。[7]

当我们试图将这种复杂性拆解开来时，难免会陷入“挑战者号”航天飞机的主流事故中。[8]尽管航天飞机的发射和操作是一项异常复杂的工作，但是很多人都相信，只要彻底核查复杂系统，将其分解开来，确定每个部件如何工作，在什么情况下会出故障，我们就应该能够理解它们。这无疑是一种过于自信的想法，其根源在于：人类的思想拥有无限潜力。人们总认为，只要足够努力，就可以完美地理解身边的一切事物，特别是人类自己所构建的那些东西。

我们观察到，这种情绪在所谓的辉格党式（Whiggish）的进步观念中体现得最为充分。正如科学作家菲利普·鲍尔（Philip Ball）所描述的那样，这种观念认为人类已经踏上了“从无知和迷信的黑暗时代走向理性时代的胜利之旅”[9]。在科学界和技术爱好者的圈子中，这个观念根深蒂固，正如历史学家伊恩·毕考克（Ian Beacock）在文章中所写：“科技行业讲述了人类‘数字时代’的崛起，那是一个辉格党式的故事：阴郁昏暗的时代行将结束，人类将迎来一个更加美好的未来，一个更幸福、更开放的世界。在那个美丽新世界中，一切都是测量好的、经过精心设计的，并达到了最高效的完美状态。”[10]当然，坚持这种观点的人还认为，效率和生产力的提高，促使人们为了实现某个令人振奋的目标，而持续关注和理解自身所构建的现代工程。这种辉格党式的观点与社会学家马克斯·韦伯（Max Weber）所描述的“现代思维模式”有关：在一个“不再拥有幻想的世界”中，人们会产生一种感觉，也就是“在原则上，我们可以通过运算来掌握所有事物”。[11]

然而，这种理解复杂性的方法如今已经失效了。是的，我们正处在一个新时代，正在建立一个新系统，而这个系统是无法从整体上被掌握的[12]，或者说，无论对谁而言，它都太复杂了。我们发现，在分析丰田汽车意外加速事故时，无论是不是专家，上述那种陈旧的思维模式都不足以应对其复杂性了。不仅如此，这些情况不只附着在我们缺失的经验中，它们已充斥了我们的整个生活。

什么是复杂系统

在这里，我们必须先花些时间讨论一下“庞杂”（complicated）和“复杂”（complex）系统的含义。[13]尽管我在文中，以及书名中使用

这两个术语时所指含义与它们在口语中的意义大致相同，但实际上，它们的准确含义和日常意义存在着一些重要区别。

请想象一下，有许多水上浮标被绑在一起，漂浮在水面。[14]当一艘船经过时，其尾流会形成一个个小波浪，从而使浮标一个个动起来。但是，每个浮标都无法单独“行动”。因为每一个浮标都通过绳索连接着不同重量和大小的其他浮标，所以任何一只浮标的行动，都会带动其他浮标产生相应的行动。这些行动还可能会引发意想不到的反馈过程，也就是说，浮标的行动会间接地影响到自身。于是，船的简单尾流，在这个复杂的浮标网络中引发了大量级联式（cascade）的行动。如果船以其他方式驶过，比如以另一种速度或角度驶过，那么浮标的行动可能会完全不同。

接下来再想象一下，人们把这些浮标从水中捞了出来，扔在了码头上。它们可能被排列得非常巧妙，却很难描述清楚，我们需要颇多文字才能讲明白它们之间的相对位置，或许还需要附上一些图表，以便日后能复制出同样的排列方法。但是，无论被排列成什么样，这里也不会有趣事发生。因为在这个精巧的网络中不存在级联效应，也不存在反馈过程。说到底，只不过是一堆可以漂浮的东西被放在了码头上。

水中的浮标形成了一个复杂的系统。那么码头上的浮标呢？它们的排列也许看上去很巧妙，但其所构成的系统最多只能说是庞杂的。任何一个系统，要想成为一个复杂的系统，仅仅包含很多组成部分是不够的，各部分之间还必须相互关联，并在“嘈杂的舞蹈”中相互作用才行。在这种情况下，我们可以观察到某些行为特征，那是复杂系统的标志：微小的变化通过网络被级联式放大，并引发反馈过程，同时敏感地依赖于系统的初始状态。这些特征，与其他一些特征一起，使系统从“庞杂的”变成了“复杂的”。

对于这种区别，还可以从另一个角度来思考：生物在活着的时候是复杂的，而死后最多只能说是庞杂的。死去的动物虽不失其复杂特性，但其内部却没有任何过程在发生：整个生命网络，比如血液循环系统、新陈代谢系统、神经系统等，都很“安静”。生命是运动和相互作用的“狂暴之流”，极其复杂精妙，在有生命的机体中，再小的变化都会被级联式放大，进而引发大量行为。另外，即使一个系统是动态的，比如浮在水中的一堆没有被连接起来的浮标，因为其内部不存在相互联系，没有反馈的可能性以及其他类似的属性，所以我们也只能称其是“庞杂的”，而非“复杂的”。

如果将技术定义为人类为特定目的而设计和构建的各种类型的系统，那么就不难注意到，当今最先进的技术几乎都是复杂的系统：动态、功能复杂、规模庞大，而且拥有近乎有机生命体般的复杂性。这些复杂的系统遍布我们周围，从汽车软件到电脑设备，再到城市基础设施。[15]那些庞大的、高度互联的软件，其规模足以与百科全书相比。据估计，微软的办公软件就有数千万行计算机程序代码。[16]

美国的公路系统有30万个交叉路口，每一个路口都有交通信号灯。[17]这是一个覆盖全美的动态交通网络，所包含的自动纠错系统先进得令人难以置信，尽管我们会因为它经常出问题而嘲笑它愚蠢得令人绝望。该系统的背后是PB级数据（1PB等于100万个GB）和复杂的概率模型。[18]随着时间的推移，法律体系也变得越来越复杂。截至2014年，美国联邦税法的总页数已经超过了74 000页。[19]这个庞大的法律网络非常复杂，拥有无数个关联点，会对纳税人产生级联式的影响，可以说几乎没有人能从整体上完全理解其功能。

在我们的生活中，复杂的技术系统无处不在，这通常是一件好事。正是在这种庞大的复杂性中，我们发现了惊人的弹性，也就是复原力。这些系统通常拥有许多功能各异的工具箱和故障保险箱，可以帮助人们完成任何“想要”完成的事情。这些系统也为我们提供了连古代皇室都无法想象的生活：免除了辛苦的重复劳动；把水电送到家中；让我们一年到头都可以生活在舒适的温度下；还能帮助我们快速地传递信息。

那么，我们该如何理解复杂系统，而这又意味着什么呢？无论是分析一个现象，还是了解一个系统，都不是非此即彼的二元选择。这种理解其实是一条内涵丰富的光谱。[20]换句话说，你可以只从整体上去理解一个系统，掌握它的全局情况，而不一定非得理解其各个组成部分的细节情况；你也可以只理解它的所有组成部分，而不太关注其整体功能；你还可以只了解各部分是如何相互关联的，或者只探究这种关联的最终效应。此外，所有的“理解”都会涉及一些具体活动：描述事物是如何运行的，在不同层面上预测其未来的行为，并在时间充足和资源丰富的情况下通过构建模型来复制它。

回到上文中有关浮标的例子。你或许只能了解到两三个被连接在一起的浮标的行为细节，甚至只能搞清楚一个浮标的行为，而无法描述所有浮标所组成的巨大网络的整体行为；或者，你能够描述浮标的整体运动，却无法预测它们各自的行为。对于软件，你或许可以很好地理解给定程序中的若干模块，例如计算圆周率值的子程序，或是对一组数字进

行排序的子程序，但你却不一定能理解它们是如何一起发挥作用的。通常，我们只能把握上述“理解”中的一部分，而非全部。

另外，“理解”也不是固定不变的，它可以通过训练来改进。从来没有下过国际象棋的人，在看到棋局的时候，可能无法将势均力敌的平局和王已经受到致命威胁的残局区别开来。而接受过一定训练的初学者和中段棋手，均能掌握下棋的基本模式，并判断棋局的基本形势和大致走向。国际象棋大师则能一眼看透当前的形势和接下来可能会出现的演变，而后评估棋局并给出可靠的走法，同时弥补潜在的弱点。经过充分训练后，棋手在看到一盘棋时，眼中所见的就不再是一排排、一列列的棋子了，而是诸如“白棋三步就可将死对方”之类的棋局。因此，学习专业知识并接受一定的训练确实可以从根本上改变我们对世界的看法和理解方式。

对于人类构建的系统而言，亦是如此。计算机代码既可能是难懂的天书，也可能是解决难题的优雅方案，这一切都取决于你到底了解些什么。不过，当我们无法完全理解它们时，就会遭到某种特定的失败或打击，也就是某种意想不到的结果。

以空中交通预警防撞系统（Traffic Alert and Collision Avoidance System, TCAS）为例，人们创建这个系统的目的是防止飞机在空中相撞。这个系统会提醒飞机驾驶员注意潜在危险，并告之如何根据规则做出应对。但是，几十年来，这个系统的规则已经变得极其复杂，复杂到全世界只有屈指可数的几个人能够真正了解它。每当有人提出一项新的系统规则时，有关方面就会通过模拟实验来测试其效果。在若干次测试中，如果表现均能达到预期，新规则就会被批准投入使用。

尽管避免飞机相撞本就是一个复杂的挑战，但是从根本上来说，我们为此而建立起来的系统已经变得太过复杂了。不但一般人无法理解，即使是专家，有时候也会对系统在某些情况下所做出的反应感到疑惑。

当结果出人意料时，就意味着我们对相关问题的理解还不够充分，因而无法搞清楚这种结果从何而来。倘若只是视频游戏中的一个错误，那倒也无伤大雅，甚至还能起到娱乐大众的效果。然而，如果是极度复杂的社会保障系统发生了出人意料的状况，我们肯定就无心娱乐了。这些复杂的系统可能是为我们提供水电的基础设施，也可能是执行金融交易的软件，还可能是防止飞机相撞的程序。在这种时候，理解是否到位就成了一个生死攸关的问题。

每个人理解世界的能力天生就有所不同。天才拥有常人难以企及的直觉，他们的跳跃性思维可能远远超出了常人所能理解的范畴。但是，人类的认知能力终归是有限的。随着时间的流逝，我们所构建的各种技术系统变得越来越复杂，系统之间的关联性也变得越来越强，越来越难以理解。无论人类有多聪明，记忆力有多强，都无济于事，因为这些系统的构建方式与人类的思维方式不一样。人类不具备同时应对数百万个组成部分及其之间的大量交互，并将所有结果都记在脑袋中的能力。我们的大脑会“严重超载”，继而宣告失败。

从“启蒙时代”到“纠缠时代”

在当今这个机器时代，非技术专家在试图理解技术时，偶尔会诉诸迷信或某种一厢情愿的观点。通常来说，在一个家庭中，总会有一个人会因计算机或某种机器无法正常运行而受到指责，在其他人看来，他只是用手碰了一下就把事情搞砸了。有时候，甚至只要他在场，人们就会认定他是导致某种技术无法正常发挥作用的原因。例如，一个大学生放暑假回到家，偏巧打印机坏了；或者父母来访，电脑鼠标就失灵了。

反之亦然。一个你无法解决的问题，别人却能轻松解决。你将某个出了故障的设备交给技术人员维修，可是一到他们手上，故障就消失得无影无踪，而当你把设备带回家后，却发现它仍然无法使用。

在缺乏专业技术知识的外行人看来，机器的内部工作原理相当神秘。如果我们只是这些系统的使用者，那么不知其详也并无大碍。当机器出现故障时，我们可能会半真半假地认为有人正在对机器“发功”，对它施加了不利影响。当然，不管有没有做出这样的假设，至少我们会很自然地相信，肯定有某个专家知道到底是怎么回事，比如“主板烧了”，于是我们便将那个神秘的捣蛋鬼因素抛之脑后。

然而，持有这种“迷信”观念的，早已不再仅限于普通人了，事实上，就连技术开发者也开始这样想了。美国计算机工程师李·费尔森施泰因（Lee　Felsenstein）讲述了一个发生在某位工程经理身上的故事：在对软件进行展示的时候，那位工程经理不得不选择回避，因为只要他在场，软件就会运行失常。软件工程师们完全搞不清楚为什么会这样，总之只要这位经理在场，事情就会变糟。[21]费尔森施泰因指出，这种无法通过运算来解释的失败“已经落入形而上学的领域”。工程师们不明就里，下意识地把思考点放在了“存在问题”这个层面上，并走上了哲学思

考之路。

当然，他们并不孤单。计算机科学家杰勒德·奥尔茨曼（Gerard Holzmann）也有同样的感受：

> 大型的、复杂的代码总会包含一些不祥的“暗码”（dark code）(1)。没有人能完全理解这种代码，它的存在也没有明确的目的。然而，想要让应用程序按照预期正常地运行，这种代码又是不可或缺的。你不想去碰它，所以会倾向于绕过它。
>
> 暗码是会逆转的。应用程序具有防止实际代码被追溯的功能，它会莫名其妙地做出一些程式之外的事。[22]

在司法领域中也存在类似的状况。根据律师兼作家菲利普·K.霍华德（Philip K. Howard）的说法，我们目前正处在这样的境况中：“现代法律的丛林已经太密集了，成了不可知之境。”[23]在技术领域，这样的例子显然更多，甚至不需要到最前沿去找。正如作家奎因·诺顿（Quinn Norton）所指出的那样，即使是功能平庸的家用台式电脑，也是“非常复杂的，复杂到地球上的任何人都不知道它在做什么、怎么做”。[24]

近几十年来，在人类一手打造的“创世界”中，异常庞杂、难以理解的事物越来越多。事实上，即使是这些系统的创造者也常常一头雾水。美国政治理论家和技术哲学家兰登·温纳（Langdon Winner）在其著作《自主的技术》（*Autonomous Technology*）中指出，英国著名小说家和政治家赫伯特·乔治·威尔斯（H. G. Wells）晚年提出的那个观点是值得相信的：“人类的思维已经无法应对它自己所创造的环境了。”[25]威尔斯早在1945年就得出了这个结论，不过，他当时讨论的主题是“人类的组织和社会本身”。近年来，这个问题变得愈加尖锐，尽管技术已经发展到当年的威尔斯完全想象不到的水平。

计算机科学家丹尼尔·希利斯（Daniel Hillis）认为，我们的世界已经从“启蒙”（enlightenment）转向了“纠缠”（entanglement），至少技术领域肯定如此：“技术已经变得如此复杂，以致我们无法完全理解它，也无法完全控制它。我们已经进入了‘纠缠时代’……每个专家都只了解难题的片段，却无法把握难题的整体。”[26]就连作为技术创造者的专家都无法完全了解技术了。

抽象的局限

在创建复杂的技术时，最强大的方法就是人们常说的“抽象”（abstraction）。从根本上说，抽象其实就是：将系统中某些不必要的组件细节隐藏起来，同时保留组件与系统进行高效交互的方式和能力。例如，我在编写计算机程序时，不必再用机器代码语言来编程，而可以使用C语言或其他类似的语言。机器代码语言是一种二进制代码，此前，每台特定的计算机都需要用这种语言来设置指令。现在，我写出来的程序不仅易于阅读，还可以转换为机器代码语言。在很多情况下，我甚至不需要知道程序将会在哪种特定的机器上运行，那些与机器有着更深层级交互的其他程序自然会“考虑”细节。换句话说，这些细节在编程时已经被我抽象化了。

这种抽象方法在技术中无处不在。在与某个界面友好的网站进行互动时，对于网站的内部技术细节，我们并不关心；在将烤面包机的插头插入某个插孔中时，我们也不需要知道电力是哪里输送过来的，或是发电厂位于什么地方。这就好比我们无须知道搜索引擎是如何给出具体结果的。只要接口或界面是合乎逻辑的、可以使用的，那么我们就只需要关注正在构建或修复的细节，而不用担心接口和界面背后的一切复杂问题。利用这种抽象方法，我们可以在一种技术的基础上构建出另一种技术，也就是直接利用他人创建的技术，而无须了解其内部细节。如果你是使用统计软件包分析数据集的财务分析师，或是使用预设代码创作有趣动画的应用程序开发者，那么你无疑已经在使用抽象方法了。

抽象给人们带来的最大好处是专业化。即使系统拥有数百万量级的交互，系统的构建者和维护者也不是必须知道它到底是如何工作的。抽象使他们只需了解自身关注的某个具体部分，而其他部分的细节会被再次抽象化。

然而在这个纠缠时代，抽象方法也可能会崩溃。事实上，这种情况出现得越来越频繁。在一个系统中，原本被设计者和构建者屏蔽的各个组成部分，正在越来越多地以意想不到的方式发生碰撞。

这一点在金融领域显得尤为突出。在当今的金融市场上，“参与者”早就不仅限于人类了，大量以各种信息为基础的计算机程序也参与了交易，而且速度比人类手动执行快了无数倍。这些计算机程序以异常复杂的方式相互联系着，并通过巨大的交易网络对决策进行级联式放大和传播。那么，它们究竟是如何做出交易决策的呢？某些计算机程序能从海量的数据中总结出有意义的参数。

结果可能会非常极端。2010年5月6日，全球金融市场出现了闪电崩

盘，股市出现了大规模的、非常迅速的巨幅震荡。[27]许多上市公司的市值都因此遭受了重创，不过不久之后又都基本重回原位。这次闪电崩盘涉及一系列交易算法和实施细则，这些算法和细则以意想不到的方式进行了交互，在短时间内便造成数十亿美元的损失。尽管很复杂，但这些系统并非处于真空地带，它们是更高层级的技术生态系统的一部分，而这个技术生态系统决定了每种证券或商品的交易时段。金融系统还会受到一系列法律法规的约束。当然，法律法规本身即是一个系统，而且是一个庞大且复杂的系统。不同法律之间存在相互依赖、相互援引的关系，且联系方式时而精确、时而混沌，令人难以琢磨。

此外，这些交易所依赖的基础设备，都建立在持续发展了数十年的技术之上。这样一来，整个系统就成了“新”与“旧”的组合：在这个系统中，古老的依靠人工喊价的实物交易与通过光纤传输完成的电子交易共存。我们在构建能够实现高效交易的计算机程序时，不仅要考虑日新月异的计算机科学、繁杂的金融工具和深不可测的法律法规等，还要深入了解物理学。因为“光在不同材料中的传输速度”这一点在交易中至关重要。因此，我们说，地球上没有人能完全理解金融世界中相互关联的所有系统，甚至没有人能完全理解其中任何一个系统。

当然，在很多情况下，系统的使用者确实只需要很好地理解系统的一小部分，甚至仅是最表层的一部分就可以了。在一家金融公司里，程序员可能只需要知道如何维护交易系统，而不需要了解计算机在物理层面上的基础设置；有的人可能只会关注某个特定的软件，这些软件可以对公司外部的消息进行过滤，然后将部分信息传输到公司内部进行运算，至于其他大部分信息，只作大致了解即可；为该公司工作的律师则需要了解与各种交易有关的法律条文，但不需要知道软件、服务器或光纤的任何细节。在这里，抽象方法给我们带来了很大的便利。

在大多数情况下，对某个事物“足够”了解，看起来已经很不错了，[28]但是，当我们构建出越来越复杂的系统之后，在系统及其子系统运行的不同层级之间就会越来越频繁地出现跨界交互现象。尤其是随着事物之间的相互联系不断加强，我们愈加难以判断原有的那些粗略的、不完整的理解是不是仍然够用。在纠缠中，事物会在不同的抽象层级上相互碰撞，以各种令人无法想象的方式进行交互。在充满交互的网络中，常常会出现被复杂性科学反复提及的“涌现”（emergence），也就是某种层次上的交互最终导致其他层次上的交互出现了预料之外的现象。涌现，在所有类型的复杂系统中都很常见。例如，昆虫的集体飞行便是一

种涌现行为。另外，在金融系统中，涌现也很常见。金融系统的运行所涉及的因素多种多样，上至全球范围内的算法交互，下至每条网络线的传输速度。要想真正搞清楚哪些细节应该被抽象化，是一个太过繁杂的问题。

当系统内部深处的某些微小细节像“微型造物主”一样崛起，并开始破坏技术系统的其他组成部分时，我们就不能只对系统进行局部理解了。当系统处于纠缠之中时，其内部各部分之间的交互会陷入混乱，以往帮助我们管理复杂性的等级结构和抽象方法，都会迅速地分崩离析。

那么，在可见的未来中，人类是否有希望找到摆脱这种混乱的办法，并让世界重新回到可管理的状态呢？或者，我们注定要带着深切的、无法言喻的恐惧感来面对这些不断增殖的系统？

在大多数人看来，不完全了解那些技术系统，不知道所在城市的基础设施细节，不理解苹果手机的硬件如何验证指纹，不清楚法律法规如何促进国际贸易，其实并没有什么关系。人们简单地认为，对于那些复杂的系统，只需要明白如何使用即可，它们的具体机制是什么则无关紧要。不过，对于一个新工具的工作原理，有人搞不懂是一回事，所有人都搞不懂又是另一回事。许多人还在自欺欺人地以为，专家们终会把我们从这种庞大的复杂性中拯救出来，因为他们理解我们所不理解的东西。但是事实并非如此，专家也不一定理解，依赖专家的时代早就一去不复返了。

我们过去所采用的用以理解这些系统的思维方式，也就是找到“挑战者号”航天飞机失事原因的那种思维方式，现在已经完全失效了。这很令人绝望。纠缠的世界并非远在天边，而是近在眼前。每个人都需要用新的思维方式去理解技术，甚至是那些我们轻松地将理解外包给专家的技术。

虽然对人类而言纠缠时代的到来是一场严峻的挑战，但我仍然满怀希望，并坚信：我们定能学会如何处理这些系统，至少在某种程度上。

但是，要想真正理解这个由人类一手创造的时代，我们就需要先退而求其次，将迫使我们陷入复杂性、阻碍我们理解复杂性的各种因素找出来。

OVERCOMPLICATED

Technology at the limits of comprehension

技术系统变得越来越复杂的主要原因是“吸积”和“交互”。随着时间的推移，系统中不断加入更多的组成部分，部分之间也增加了越来越多的连接。“必须处理的例外情况”和“普遍的稀有事物”也让技术系统变得愈加复杂。

02

复杂系统形成的 4 个原因

技术系统变得越来越复杂的主要原因是“吸积”和“交互”。随着时间的推移，系统中不断加入更多的组成部分，部分之间也增加了越来越多的连接。“必须处理的例外情况”和“普遍的稀有事物”也让技术系统变得愈加复杂。

要想使用互联网，我们就必须忍受或是间接地忍受杂乱无章、东拼西凑的网络世界。真是一团糟！互联网的第一步是在20世纪60年代迈出的。[1]当时，有人创造出了一个巧妙的设计，使人们能够在不同的地方通过计算机传递信息包。于是，小网络被相互关联起来，构成了大网络。同时，为了高效地传输信息，人们还开发出了各种各样的协议。

到了今天，互联网的应用已与其最初出现时大不相同，以安全问题为例。互联网原本是研究者以沟通为目的开发出来的系统，对于要求高效和安全的大规模商业交易而言，并不理想。为了弥补这种缺陷，促进商业交易，人们在原有的互联网基础设置上开发出了很多不同的机制，包括加密和解密信息的方法，以数字方式转移资金的规则，等等。令人高兴的是，这个系统确实变得有用了。然而，在网站的用户界面背后，其实潜藏着一个奇怪且复杂的结构。有时候，作为用户的我们也能直接窥见些许混乱的迹象，例如，网页上弹出的安全证书警告。是的，很多事物都是这样，有用，但是远远称不上优雅。

与此类似，开发网站所用的HTML语言，在设计之初，也并非是为了服务诸如谷歌在线办公软件（Google Docs）这类基于全球广域网（Web）的交互式应用程序。现在，这类应用程序已经投入运行，但人们仍然在付出代价：我们必须在一个简单的系统上建造一栋非常宏伟的巴洛克式建筑。倘若你想一瞥这种复杂性的一角，只需查看一下谷歌主页的源代码就足够了。虽然我们在浏览器中看到的网页既简洁又优雅，但是潜藏在这个表象下的东西却数不胜数。我上一次查看谷歌主页的源代码时，其字符数已超过了10万个，如果完整地打印出来，将超过50页纸！[2]

再比如电子邮件。从表面上看，这是一个相对简单的应用程序，已经发展了好几十年。在其古老的原始结构的基础上，消息线程等各种新功能层出不穷。不过，网络杂志《页岩》（*Slate*）的互动编辑克里斯·柯克（Chris Kirk）在尝试构建自己的电子邮件客户端程序之后指出："虽然电子邮件的软件时有创新，但这些创新都建立在过时的系统之上，如同狡猾的平衡术，有时甚至相当随意，比如，将电子邮件恢复为最初形式，或是将它改头换面。"[3]

在计算机学和工程学中，有一个术语kluge，指的是拼凑起来的系统，也就是将许多不同的东西混合在一起，以求解决问题的系统。这种系统肯定是不精致、不优雅的，而且很多时候庞杂得毫无必要。虽然这种拼凑起来的系统是有效的，但远远称不上完美。有些东西的第一代设计可能是相当优雅的，但是随着时间的推移，它们的结构变得越来越复杂，最终变成了杂乱的鲁布·戈德堡（Rube Goldberg）式(2)的应急之物。

包括互联网在内，在每一个技术领域中，都存在着这类拼凑起来的系统，例如交通网络和医疗设施。以家庭娱乐系统为例，它们虽然有用，但需要同时使用好几个遥控器，以及一大团乱麻般的电线、信号线和数据线。

美国的法律体系也是个拼凑起来的系统。这是一个为了达到特定目的而被构建出来的技术系统，远远称不上优雅。就像计算机代码是操作软件的书面描述一样，法律法规本质上也是技术代码的书面体现。[4]

毫无疑问，《美利坚合众国宪法》是一部非常优雅的文件，只用了寥寥数页，就为代议制民主奠定了坚实的基础。当然，宪法的确立并不是故事的结局。对联邦法律具有指导意义的《美国法典》是在宪法的框架内发展起来的。这些法律致力于阐明宪法的一般原则，以及对各种具

体情况的处理准则和方法。例如，《美利坚合众国宪法》只用一句话规定了国会有权建立公共邮政服务机构，而在《美国法典》中，有关这个政府职能的阐述多达500余页。[5]此外，美国联邦邮政法规还规定了从邮政局的职位设置到邮政资费等各方面的所有细节。总而言之，《美国法典》比《美利坚合众国宪法》要复杂得多。[6]事实上，《美国法典》的规模和互联性仍在不断增加，时至今日，其总字数已经超过了2 200万，内部各章节之间的关联点也已超过了8万个。

其实，无论在什么地方，我们都可以观察到，随着时间的推移，各种系统的复杂性都出现了大规模增长。我们还发现，一般来说，当一个复杂的系统庞大到一定程度时，无论其具体形式如何，都会变成一个拼凑起来的系统。莱特兄弟于1903年制造的飞机是简约主义的杰作[7]，只有很少的几个部件，载人后总重量仅为340千克；而到了今天，制造一架波音747-400飞机[8]需要用掉67 000千克铝材，600万个独立部件和275千米管线[9]。这是个普遍现象，在过去的200年里，我们制造出来的机器所包含的零件数量一直在大幅增加。

那么，对于影响着我们生活方方面面的，现代技术系统中的软件来说，情况又如何呢？衡量软件复杂性的常用指标之一是程序代码的行数。据估计，微软操作系统的源代码行数近10年增长了10倍。[10]Photoshop的源代码行数在过去20年里爆炸式增长，几乎是1990年的40倍。[11]

在电话通讯系统中，类似的情况同样存在，随之而来的还有巨大的复杂性。20世纪20年代，美国的电话通讯系统已经拥有了大约480万千米的收费线路和大约1 700万部电话。[12]要知道，就当时而言，电话才刚刚问世几十年；而时至今日，相关的技术生态系统已遍布全美。

所有这些系统都是为了实现某个特定功能，由一代又一代的专家设计、构建出来的。有人可能会认为，如果这些系统的设计是合理的，那么它们理应合乎逻辑、优雅，甚至简洁，也理应易于说明、易于修复。然而，尽管我们已尽了最大努力，但技术系统还是变得越来越复杂、越来越庞大。这绝非偶然，技术发展过程中某些固有的力量，使我们在“复杂性”中越陷越深。和万有引力之类的物理定律截然不同，这些力量强大到能让系统变得越来越复杂。不管时代如何变迁，它们总能压制住人们对简单的渴求，以致在人们心中，它们已如物理规律般不可抗拒。可是，为什么会这样呢？

在本章中，我将详细分析一些会使系统变得日益复杂的因素。从表

面上看，这些因素完全合理，它们所引发的每一个变化都能使技术更加适应不断变化的环境，要么有助于系统继续在新环境中正常运行，要么增加了系统的实用性。然而，这些因素最终会使原本优雅的解决方案变成杂乱无章的、拼凑起来的系统。无论付出多大努力也无法避免这样的结果，无法阻止技术复杂性的不断增长。最终，我们生活的方方面面都会受到影响。

各种技术系统皆会随着时间的推移而越来越复杂，究其首要原因，也是最显而易见的原因是系统内部存在着双重力量：吸积（accretion）和交互，也就是说，随着时间的推移，系统的组成部分越来越多，同时组成部分之间的关联也越来越多。

原因1：吸积

在2000年1月1日到来前的几年里，许多工程师都在研究如何解决“千年虫”问题。简而言之，若一个软件在存储年份时使用的是两位数，而不是四位数，那么当2000年到来时，这个软件就会“认为”那是1900年，从而引发诸多问题。恐怕没有任何人、任何机构会比美国联邦航空管理局（FAA）更担心“千年虫”问题了。如果空中交通管制系统在新千年到来时出现故障，那将会导致巨大的灾难。因此，美国联邦航空管理局检查了计算机系统，并对系统进行了测试：如果系统认为那是1900年，会出现什么情况？[13]

在测试过程中，他们发现交通管制系统中的IBM 3083计算机问题特别棘手。作为技术人员代表，美国联邦航空管理局的工会主席指出主要问题在于：“IBM公司只有两个人知道这种型号的计算机的微代码，但是他们都退休了。”[14]但真正的原因是，IBM 3083属于大型机，从20世纪80年代开始发售，而且所用的系统软件早在发售前几年就已投入使用。这就意味着，到了20世纪90年代后期，这个为全美的飞机制定航线的计算机系统所使用的代码几乎已无人可识。

这还算不上令人震惊。一直以来，许多大型系统的基础都是较小和较陈旧的系统。只要这些系统能够继续平稳地运行下去，就不会有人在意那些旧东西上面到底堆了多少新东西，系统中到底累积了多少次第加入的片段。据一位消息人士透露，直到2007年，美国国税局所使用的报税系统依旧是20世纪60年代早期，也就是肯尼迪政府时期开发的系统。[15]美国国税局所使用的另一个系统则始建于20世纪70年代，并于1985

年进行了大修。与此类似，美国的航天飞机在执行最后一次航天任务时，[16]所用的平台是由5台IBM计算机组成的，其计算能力甚至还比不上今天的一部普通智能手机。然而，这些软件和技术仍然在使用着。

1975年出版的《人月神话》（*The Mythical Man-Month*）一书中，计算机科学家小弗雷德里克·布鲁克斯（Frederick P. Brooks Jr.）讨论了有关软件设计和项目编程的管理问题。在这本书中，他引用了业内的一句俗语："每次加一点，每次加一点，最后就有了一大堆。"[17]每一个独立设计，无论是为了修复，还是用来提供新功能，看上去都不过是一次独立的选择，而且都很合理：要么解决了问题，要么为用户创造了新的令人兴奋的功能。然而，日积月累下，它们终会变成"一大堆"。从交通运输业到能源业，再到农业，我们都可以从中清楚地看到，但凡是大型的技术系统，就定会发生这种情况。

举例来说，一大堆石头不一定会成为问题，即使它们看上去可能很难处理、过于零乱，但却不一定深奥难解。真正的问题是，当我们创造的"一大堆"引发出意料之外的状况时，我们就会遭遇"雪崩"。不幸的是，我们不断地往技术系统中加入一个又一个片段，"雪崩"在所难免。"一大堆"不仅变得更大，而且变得更难琢磨了。

我记得我最早是在一些讨论"行星系如何形成"的文章中看到"吸积"这个术语的。行星系由一团旋转的尘埃和气体凝聚而成，这种星星点点的累积过程，就是"吸积"的过程。[18]这个用来描述行星系如何形成的概念由来已久，而在技术的增长过程中也是类似的吸积作用。

吸积过程的结果之一就是形成了人们常说的遗留代码（legacy code）或遗留系统（legacy systems），即过时的机器和技术，也就是开发出来之后使用至今的机器和技术，譬如美国国税局所使用的报税系统。这种老旧的系统并不罕见，甚至可以说相当常见。[19]它们经历了多年的吸积过程，成了拼凑起来的系统，从科学模型到城市基础设施，几乎无处不在。例如，在城市的排水系统中，既有服役超过百年的旧管道，也有刚刚埋设好的新管道。

就计算机领域而言，很多技术系统都依赖于已停产的旧型号计算机系统，而且程序代码也是用早已退役的编程语言所写。例如，许多科学软件现在都成了遗留系统，它们一般都是用Fortran语言编写的，那是一种功能强大但早已过时的编程语言。随着技术的高速发展，如今我们再来看Fortran语言编写的计算机代码，就好像是在看中古时代的英语。

在这里，不妨引用《全球概览》（*Whole Earth Catalog*）的创始人斯图尔特·布兰德（Stewart Brand）在《万年钟传奇》（*The Clock of the Long Now*）一书中的说法："通常，这些已经过时的遗留系统，在过去的许多年里，一直都扮演着十分重要的角色，若是替换掉它们，必定会'伤筋动骨'，恐怕没有人能够承担这样的后果。此外，它们也是无法被完全修复的，一是因为问题过于复杂，二是因为没有人完全了解整个系统。"

那么，当我们面对一个仍在缓慢增大、小故障不断的遗留系统时，又该怎么办呢？我们只能小心翼翼地对待它，因为设计它的人可能早就"杳无音讯"了。这个遗留系统可能已经完全嵌入了其他系统，彻底移除它的后果，可能远比容忍它的小故障更加糟糕。这种系统实在难以处理，有时甚至会被称为"恐怖爬行兽"（crawling horrors），也就是美国恐怖、科幻小说作家霍华德·菲利普斯·洛夫克拉夫特（H. P. Lovecraft）小说中那种无法形容的怪物。[20]

司法系统条文，也有类似的情况。在法律体系中，随着时间的推移，法律条文会被修正或修改。人们会根据不断变化的环境，对法律条文做出调整，但总会留下一些数十年前制定的法律条文。例如，互联网流量管理的相关法规就源于1934年通过的一项法律。随着时间的推移，法律体系也在不断地吸积，最终成了一个拼凑起来的系统。它可以发挥作用，但是远远谈不上优雅。

以税法为例。事实上，美国的税法早已复杂不堪了，立法者也早就承认了这个事实。税表的使用说明已经从1940年的两页，增加到了2013年的200多页。[21]鉴于此，如果你在申报纳税时，因为法律法规太过复杂而出了错，那么最高法院会裁定你无罪，因为你在错误报税这件事上并没有"主观故意"。[22]尽管如此，就法律体系本身而言，在现有税法上修修补补，要比重新制定更容易，也更有效。

另外，我们也要考虑一下诸如美国环境保护署之类的政府部门和机构所颁布的法规的总体增长趋势。事实上，只要看一看《美国联邦法规》的总页数就足够了。[23]《美国联邦法规》收录了各行政机构颁布的各种法规，在过去的50年里，其页数从不到25 000页增加到了超过165 000页。

我们还观察到，行政人员和行政机构的数量也出现了类似的增长趋势。20世纪50年代，《经济学人》杂志中有一篇文章提到了帕金森定律（Parkinson's Law），并定量地描述了行政人员数量的增长规律。[24]虽

然这篇文章的观点还不太完善，但帕金森定律毕竟有数据的支持，其结论到今天仍基本成立：政府部门的行政人员数量以每年5%至6%的速度在增长。毫无疑问，随着行政机构的规模越来越大，机构的管理问题只会越来越复杂。

事实上，软件界已经将吸积和积累奉为普适规则。[25]就发展而言，软件系统的规模势必会与日俱增，除非有人积极地尝试简化它们。

那么，为什么我们不能对复杂的系统进行定期清理并从头开始呢？这和实际操作有关。例如，软件未能按时重写完，索性就在推出新版本之前先发布一个补丁。[26]我在想，只要愿意花上很多年的时间，那么微软内部的任何一位软件专家都可以重写Word的全部代码。当时间、精力和金钱都有限，且不得不进行权衡时，[27]我们通常会选择对系统进行修改，让它“足够好”就行了。这意味着，我们需要不断对系统进行调试和修正，就像立法者对美国法律体系所做的那样，也就是说，我们需要在以往的基础上加上一层又一层的东西。我们的城市拥有一个多世纪之前埋设的燃气管道[28]、20世纪30年代建成的运输网络[29]，以及废弃的地铁站。它们都隐藏在城市的地下。

不过，更多的时候，我们之所以放弃从头开始的想法，是因为那样做不仅太困难，而且太危险。没有人能够完全理解一个系统所依赖的所有旧的组成部分的全部作用，所以重新设计一个未经检验的系统不仅是愚蠢的行为，更是危险的行为。试想一下，一个几十年前设计好的、非常复杂的银行软件系统正慢慢适应着各种先进的技术，无论是新型的计算机，还是新的操作系统，或是无处不在的互联网。虽然这个系统的核心基础并不适用当今时代，但它们已经嵌入得太深，以致无法删除。总而言之，我们必须接受这样一个普遍规则，那就是：无论何种技术系统，终将变得日益复杂。[30]

但是，当仔细观察技术系统中的遗留代码时，无论是在一个软件中，还是在一个法律体系中，我们都会发现，真正的复杂性绝不仅仅只体现在日益扩大的系统规模上。毕竟，只有和另一个因素结合起来，吸积才能使技术系统变得复杂，这个因素就是交互。

原因2：交互

现在，几乎所有的学生都会在师长的引导下，学习一些编程技术。是的，不管是教育工作者还是技术专家，都告诉我们计算机编程就是未

来。计算机操纵了周遭的一切，比如汽车和微波炉。专家们当然没有说错。不仅如此，编程还可以为你带来一种结构化的思维模式，以及提示你关注某种技术系统的实际功能。当我说起某个新的应用程序能够做某件事时，如果你参与了编程，哪怕只是一点点，你也会比一般人更清楚我的说法是否合理。

如果你以前写过代码，那么你应该清楚计算机程序是如何抵制简化工作的。随着计算机程序变得越来越庞大、越来越复杂，代码也变得越来越复杂，而且绝大部分代码的聚合方式都是令人费解的。为了确保大型程序仍可控，我们开发出了各种各样的技术方法，例如版本控制、错误跟踪，以及跨团队沟通工具等，但是这些方法通常都只是“战斗失败”后的亡羊补牢之术。不仅软件代码本身在不断吸积，每个组成部分也在越来越频繁地与其他组成部分进行交互。与存在于“真空”中的全新项目有所不同，任何一个计算机程序都是一个大规模的互联系统：不仅会作用于自身，还会与其他程序相互作用。我们在旧代码中一次又一次地加入新代码，并以出其不意的新方式去使用它们，以此将各个层次拼接在一起。

交互过程，包括许多意外交互在内，有时会因编程语言本身的原因而加剧。即使是刚入门的程序员也知道，GOTO语句会带来麻烦。在BASIC编程语言中，通过设置GOTO语句，可以让程序轻松地从某一行跳转到另一行。换句话说，如果代码中包含了GOTO语句，那么程序就可以轻松地从某一行指向另一行，从代码中的一个点跳转到另一个点。如此容易便能实现跳转，难怪有人会说，对于那些想要对自己的计算思维进行测试的文科生而言，GOTO语句无疑是“天赐良品”。[31]

在一个小程序中，使用GOTO语句实现跳转，既容易又无伤大雅。但是，随着程序变得越来越大，GOTO语句最终会将代码绑定到某些庞杂的节点上，但即使是最熟练的程序员也无法解开这些庞杂的节点。最终，你得到的将是所谓的“意大利面条”式的代码，因为所有东西都纠缠在了一起，既难解开，又难理解。在这种情况下，要想搞清楚程序指令执行的次序几乎是不可能的，这也是为什么这样的计算机程序特别容易出现令人意想不到和无法理解的行为。

此类简化命令即使只是在独立的、较小的环境下发挥作用，也有可能通过某种方式脱离当初设定的目标。如果使用的频率和方式超出了预设范围，系统及其组成部分就会出现诸多问题，而这些问题比当初设想的要庞杂得多。GOTO语句就是一个突出的例子，它从一个美妙的超快

捷工具变成了既不优雅，又实为“有害”的东西。[32]

为了更好地理解并强化系统秩序，专家们想出了很多方法，包括使用更加复杂的计算机语言。[33]在大多数由专家构建的新系统中，意大利面条式的代码已成为过去式。然而，由于互联的易发性和各种层次上的不断吸积，交互作用正在持续增多。于是，高度互联的系统动态，即信息的流动和各部分之间的交互，也就变得异常复杂和不可预测。仍以丰田汽车软件为例，因为存在交互，丰田汽车软件中的大量代码已无法被测试了。[34]

在我们构建的其他类型的技术系统中，互联性同样在不断升级。例如，在法律体系中，由于每项新的法律法规都与以前的法律法规相互关联，因此我们很难预测单项法律条文的效果。

菲利普·K.霍华德曾仔细分析过贝永大桥（Bayonne Bridge）这个案例。[35]贝永大桥是连接纽约州和新泽西州的主要通道之一。由于这座百年大桥的桥梁实在太低，以致那些前往纽瓦克港的现代集装箱船无法顺利通过。纽瓦克港是一个重要的商业中心。那么应该怎么办呢？在人们提出的各种解决方案中，有一个方案是这样的：对这座桥进行改造，适度提高桥梁高度。这应该是成本最低的方案了，也是在2009年脱颖而出的方案。但是，改造工程拖了很多年都未能启动，因为人们难以应对吸积和交互的综合作用。与大桥改造工程有关的规章制度总共涉及19个政府部门的47份许可文件，从环境影响评估报告到历史影响评估报告，不一而足。[36]其他地方也出现过诸多类似情况，一些公共项目需要10年左右的时间才会获得批准，因为相关的规则和流程冗长繁多。[37]正如霍华德所说，这种情况在很多时候甚至是致命的，比如，老化腐朽的基础设施如果未能得到及时修缮，就有可能夺走许多人的生命。

研究员迈克尔·曼德尔（Michael Mandel）和黛安娜·卡鲁（Diana Carew）就职于位于华盛顿特区的进步政策研究所（Progressive Policy Institute）总部。他们将规则体系的增长称为“监管积累”（regulatory accumulation），即随着时间的推移，规则会变得越来越多。[38]换句话说，每一条法律法规都是合理的，但当它们被放到一起时，就有可能会因为相互作用而变得异常“软弱”，甚至可能以令人惊讶和意想不到的方式产生冲突。

我们不仅越来越多地将某项技术的各个组成部分关联起来，而且还越来越多地将不同的软件和技术关联起来。后者是一种高阶互联模式，也就是互操作性（interoperability）。[39]

让各项技术互通，也就是让不同系统进行交互，相互传递信息通常是一件好事。例如，因特网之所以拥有如此强大的功能，就是因为其连接的机器数量极其庞大，而且可以在无数机器之间传递信息。当你问Siri“世界总人口是多少”时，你的苹果手机会通过Wolfram Alpha服务获取到答案，然后回答你；当你使用谷歌地图时，它会告诉你利用“优步”去往目的地可能要花多少钱。这些都是互操作性的实例。但是不要忘记，让不同系统互通的同时，我们不得不去面对陷入复杂世界的巨大风险。我们现在不仅建成了互联网络，比如由不同计算机和设备组成的互联网，而且还建成了拥有众多子系统的大型互联系统。

除了互操作性之外，不同类型的技术之间还会产生相互依赖性，例如互联网与电网之间的相互依赖性。[40]研究者在研究了多种类型的系统，并了解了它们的优缺点后指出，某些系统在多种条件下均可能会出现故障或崩溃。例如，某个规模相当小的电网出现了故障，继而引发了无法收拾的级联效应。对于这种风险，有一种观点是：将技术系统之间的相互联系切断。然而，这种想法的可操作性几乎为零。互联系统的构建成本其实很低：在当今这个充满互操作性的时代，工程师和设计人员都在有意为各个系统创建接口，因此，不同的系统可以很容易被关联起来。

我们在构建新事物时，通常都会在故障成本和构建成本之间进行权衡。我们需要知道，如果出了故障，失败的成本会有多大。[41]如果Word崩溃了，那么尚未保存下来的东西将会丢失。尽管没有人希望看到这样的结果，但这个故障的成本的确是相对较低的。如果电网出了故障，并导致美国很多地区停电，那么故障的成本就极其高昂了。例如，在2003年，美国东北部的大停电对5 000万人的生活和工作造成了影响，并导致11人丧生，直接损失估计高达60亿美元。[42]

每一项故障成本都应该拿来与系统的构建成本进行比较。纵观历史，我们所构建的系统越重要，构建成本就越高。例如，构建银行系统的基础设施所耗费的资源，比编写一个聊天程序多得多。因此，我们必须确保那些昂贵的系统不易发生故障，而这又意味着需要增加构建成本。换句话说，在极高的构建成本面前，通过大量检查和测试来降低故障成本的做法，变得至关重要。

在一个相当长的时期内，这种方法一直行之有效。因为对构建成本的重视程度超过了对故障成本的重视程度，所以我们所依赖的所有重要的社会保障系统，都是花费大量资源构建起来的。[43]然而，现在事情已

经发生了变化。出于各种原因，例如，我们可以找到现成的工具和组件，同时“云”上面也存有很多可用资源，所以构建成本已大幅下降。创立技术公司已不再需要太多启动资金：你可以快速地设计和生产出复杂的工具，并通过市场进行测试，而你为此所付出的成本并不大。

与此同时，与互联有关的故障成本也已出现持续上升的趋势。尽管系统关联技术相当简单，而且成本很低，但是这种互联系统的故障成本却非常巨大。当我们把数字地图软件与出行指导软件关联到一起时，哪怕只是一个很小的错误也有可能导致一场灾难。例如，苹果地图在首次发布时就曾将超市错标为医院。

在当今时代，利用互联网信息人工合成微生物并不是一件不可思议的事情，但正因如此，爆发生物灾难的风险也比以往任何一个时代都高得多。例如，已经有实验室通过使用电子邮购的生物原料合成出了脊髓灰质炎病毒。现在，有不少初创公司正在努力实现生物学实验的远程操作。不难想象，在这个日益自动化的世界中，软件合成的生物因子[45]完全有可能会在不经意间被释放。当构建成本持续暴跌，而故障成本直线上升时，我们便进入了一个无比复杂的技术领域。是时候停下来想一想了。

一般来说，随着系统内部及系统之间的交互增多，包括拥有子系统的大型系统在内的所有系统的复杂性都会增加。有人认为，互联性的不断提高，体现了技术的基本要求。[46]技术终究会产生交互和聚合，并在这种情况出现时进一步推动我们走向复杂化。

尽管从开始构建大型技术系统的第一天起，这些趋势就一直存在，但是近年来，它们变得愈发强大了。正如我在本书导论中所提到的那样，计算机科学家艾兹格·迪科斯彻对当下的大型系统，特别是计算机系统的激进新颖性所进行的分析发人深省。早在1988年，迪科斯彻就已指出，计算机程序的设计需要克服大量规模上的差异，当然，在计算机问世之前，没有人需要去处理这样的事情。[47]以智能导航系统为例，迪科斯彻很好地解释了其跨度极大的层级结构。从程序中的1个比特，到机器存储空间里的几百兆字节，的确是从非常小到非常大的跨越。这种跨越涉及近10亿次的跳转，极端的规模变化不但超乎想象，而且史无前例。事实上，这一切已经，并将继续走向极端化，因为日常应用技术的普通用户，现在已经熟悉了千兆或万兆这样的前缀，而这些前缀又意味着，我们要对庞大的规模差异负责，而这种庞大系统的边界已近乎天文意义上的边界了。

在过去的短短几十年间，大型系统已变得异常庞大且错综复杂，用迪科斯彻的话来说即是“概念层级的深度，绝非人类心智曾需面对的任何事物可比”。

不仅如此，即使我们有能力阻止系统的吸积和交互，也还需面对另一个会使系统变得日益复杂的因素，而那将是更加难以解决的问题。

原因3：必须处理的例外情况

假设你想创建一个日历程序。看上去很简单，是不是？从很多方面来看，的确如此。要想计算一年的天数，相对来说并不算困难。一般来说，一年有365天，遇到闰年就在2月底加上一天。怎样计算闰年？也不难。只需要看一下年份即可。如果年份可以被4整除，但无法被100整除，或是能被400整除，那么这年便是闰年。

也许你还希望这个应用程序能够处理时区问题。这应该也不算难。只需利用全球定位系统（GPS）中的地理坐标来确定所在位置的时区即可。你还可以创建一个列表，将各州与时区对应起来。当然，时区并不是沿着州界划分的。现在你需要用更具体的、分辨率更高的信息来处理小区域。此外，你还不能忘记，亚利桑那州的大部分地区都是不使用夏令时的，它把自己置于一个“特殊的时空”当中了。

是不是还想让日历程序包含节假日的相关信息？当然，节假日通常都有明确的时间安排，至少大部分节假日都是如此，把它们加上去应该不会太困难。感恩节是11月的第4个周四，美国退伍军人节被定在每年的11月11日。那么逾越节（Passover）呢？看来，我们还需要将这个日历程序与另一个基于希伯来历的日历程序整合到一起。因为逾越节是从希伯来历一月的第15夜伊始的。因此，我们需要远超预期的更多信息。

那么，你是否还希望这个日历程序包含其他时间段，并且在过去时段和未来时段上都表现得准确无误呢？如果我们回到19世纪，在那个时代，由于标准化时区制尚未被推行，所以各个城镇都有自己的时间制度，而这些信息都需要“硬编码”（hard-coded）[(3)]到我们的应用程序中。与此类似，在过去的几个世纪里，虽然全球许多国家和地区都已放弃使用罗马儒略历（Julian calendar），转而使用格里高利历（Gregorian calendar），但是各国以及各地区所采用的具体时间制度并不全然相同。例如，俄罗斯“十月革命”的纪念日之所以是在11月，是因为在革命发生时，俄罗斯仍在使用罗马儒略历，其日期与西方许多地

区所使用的格里高利历相差一周多，而“十月革命”发生在儒略历的10月底。如果希望日历程序准确且详尽，那么就应该将此类信息也植入进去。

这个过程还将持续下去。

以这种方式构建起来的系统最终会变得复杂无比，因为它所要反映的事物，本身就是复杂的。[48]通过一个简单的模型来处理绝大多数复杂性，是相对直接的方法。比如，我们知道了一年有365天或366天，就可以通过简单的运算来确定某一年到底有多少天。但是，如果你对准确性有要求，无论是想确保永远不会错过任何一个约会，还是想构建一辆既不会迷路也不会撞伤人的自动驾驶汽车，事情就会变得非常复杂。[49]

这种复杂的情况就是必须处理的例外情况，也就是所谓的“边界情况”（edge case），若不处理，技术系统就会出现漏洞。[50]边界情况各种各样，从闰年问题，到如何编写数据库软件来处理特殊的人名，比如人名中带有特殊符号的情况。我们不能说边界情况是普遍现象，但它们确实经常出现，所以我们必须加以识别和管控。但是与此同时，技术的简单性也就渐渐消失了。边界情况使技术变得复杂了。这一点在科学模型中尤为突显，科学模型也是一种技术，也会随着时间的推移而发生变化。接下来，我们就以社会科学中的语言学为例展开讨论。

原因4：普遍的稀有事物

在中学时期，英文老师教会了我语法，因此我很早就知道，不规则动词to be的两个单词是必须连在一起用的；我还记住了很多介词，并学会了造句。在那个时候，根据语法规则来分解句子，将句子的修饰成分去除，剥离出它的逻辑骨架，是一件很有趣的事情。你可以将语言简化到原子量级，如名词、动词和形容词，然后再来看它们是如何关联在一起的。

尽管语言无法用方程式来表达，但是语法确实拥有一种独特的、有秩序的美感。然而，构建一个语言处理系统并不是一件容易的事。任何一种语言都有很多习惯性表达，而且其内涵往往比我们想象的要丰富得多、“狡猾”得多。因为语言具有非正式性，所以使用者在面对作为规则集合的语法时，多半只会深表认同，而不会严格遵从。所有这些都属于同一类边界情况，它使下面这个简单规则无法成立：每个句子都必定是“主语－谓语－直接宾语”这种结构的变形。为了更好地理解语言中的

边界情况，我们现在来讨论一下所谓的罕用语（hapax legomena）。罕用语可谓“普遍的稀有事物”（common rarities）。

请问你以前用过snowcrie这个单词吗？我想你应该没有用过。事实上，snowcrie这个词是无意义的。据我所知，它可能是一个错词。根据《牛津英语词典》的解释，snowcrie这个词曾经出现在1402年的一首诗中：“Not in Goddis gospel, but in Sathanas pistile, wher of sorowe and of snowcrie noon is to seken.”。有学者认为它应该是一个错词，[51]正确的那个词可能是sorcerie，意为巫师。

不管有没有意义，snowcrie这个词就是所谓的罕用语，或者说“只用过一次的词”。这个词在《牛津英语词典》的语料库中只出现过一次。语料库是大量的、通常是完整文本的合集，例如某种语言的全部文本，或某个时期的所有文本。《牛津英语词典》的语料库即是编写者可以使用的所有英语文本。不过，语料库的文本体量并不一定非常大。在莎士比亚文集这个语料库，也就是莎士比亚的全部著作中，经常会碰到一些罕用语，比如honorificabilitudinitatibus，含义可能就是“荣誉”（of honor）。

当一个语料库拥有某种语言的全部，或近乎全部的文本时，罕用语就会变得让人头疼，比如《希伯来圣经》中的希伯来语，人们对它们的意义知之甚少。但是，罕用语并不是离奇的统计错误。它们不仅比想象中更加普遍，而且与语言学中特定的数学规则有关。语言中不同词汇的使用频率可以用幂律（power law）(4)中的长尾分布来进行描述。[52]长尾分布与用来描述人类身高的钟形曲线（bell curve）(5)，即正态分布，有所不同。在长尾分布中，有一些值会延伸到更加深远的区域，以便容纳普通词汇，比如the，或一些极其罕见的词汇，比如flother。

一般来说，语料库中有近一半的单词都只出现过一次，也就是说，有一半的单词都属于罕用语。这些词就是长尾中“长”的部分。[53]因此，虽然你遇到某个特定罕用语的概率很低，但是你遇到这类词的概率却相当高。在这里，我们不妨用看电影来做下类比。相信并没有太多人看过那部大名鼎鼎的经典电影《天生爱神》（*The Adventures of Buckaroo Banzai Across the 8th Dimension*），但是看过至少一部经典科幻电影的人却大有人在。

因此，作为一个整体类别，罕用语是非常重要的。它们深深地渗透在我们的语言之中。当我们试图编写一个计算机程序来模拟语言时，可能会将罕用语，或罕见的语法结构抽象为异常值。但是，作为一个类

别，而非一个单词，罕用语在语言中所占的比例其实是相当大的。将它们抽象化会导致模型的严重缺失，从而使程序变得不完整。为了避免“遗漏”，[54]我们需要建构出可以处理例外情况和边界情况的复杂模型。在这个问题上，谷歌公司研发主管彼得·诺维格（Peter Norvig）的话可谓一语中的：“形成一种语言的，并不是那种可以用几个参数来代表的永恒的理想模型，而是复杂过程中的偶然结果。”[55]

因此，计算机语言学家应该考虑边界情况，并尝试着针对这个复杂系统构建一个稳健的、丰富的技术模型。在这里，复杂系统指的就是语言。那么，他们最终会得到什么呢？毫无疑问，他们将会得到一个复杂的技术系统。

和语言有关的计算机模型必定具有复杂性。要说明这一点，只需举一个例子就够了：计算机是如何将一种语言翻译成另一种语言的。关于计算机的翻译功能，有一个流传已久但未经证实的故事。[56]在冷战期间，科学家们就已开始研究英俄语互译的运算方法了。在测试计算机的翻译程序时，他们选择了一个含义相当微妙的句子“The spirit is willing, but the flesh is weak.”（灵固有所愿，肉却软弱不堪）。他们通过计算机将这句话翻译成俄语，然后再次翻译成英语，最终得到的是“The whiskey is strong, but the meat is terrible.”（威士忌很有劲，但是肉却很难吃）。

显然，通过运算来实现机器翻译功能并非易事。谷歌翻译虽然有趣，但结果却可能不够准确。不过，专家在这方面已经取得了很大的进步。

那么，机器翻译专家使用的是哪些技术呢？[57]早期的一种方法是利用语言的结构化语法，进行模型搭建。计算机语言学家将每种语言的属性硬编码为软件，然后让计算机根据语法规则进行翻译。这种方法可以处理相对简单的句子，但无法应对日常语言的多样性。例如，一个用来处理“直接不定式”的规则，不一定能够处理“分离不定式”，即无法处理“To boldly go where no one has gone before.”（勇敢地进入前人未曾涉足之地）这样的句子。另外，不定式的用法还具有一定的地域性，比如，匹兹堡人很喜欢省略掉to be，直接说“The car needs washed.”（这车该洗了）。很显然，面对这种形式灵活的“方言”，语法规则将束手无策。

事实上，依赖上述语法模型的机器翻译程序是不可能给出准确结果的。语法规则看上去既优雅又简洁，但无法应对文本翻译过程中所需处

理的，复杂且古怪的语言现象。简而言之，边界情况实在太多了。为了填补这道鸿沟，机器翻译专家引入了机器学习领域的多种统计方法，他们让计算机先摄取大量已经翻译好的文本，然后基于一组算法翻译新文本。这样一来，计算机就不用理解句子的含义，也不必解析句子的语法结构了。举例来说，对于复数问题，我们不再需要创建复数的语法规则，规定将后缀“-s”加在单词末尾，就能够使之变成复数形式；我们只需让机器知道，“-s”这个后缀在99.9%的情况下意味着创建了一个复数形式的单词，而在剩余0.1%的情况下并非如此。[58]例如sheep和deer，它们的单复数形式相同；此外，还有一些单词的复数形式是不规则的，如men、feet和kine。对于语言系统中的其他例外情况，也都可以采用这种运算方法。

尽管摆脱混沌就能迎来秩序，但这不可能没有代价。最终得到的最有效的翻译程序肯定不是一个简单的模型，而会是一个拥有大量参数的庞大的计算机系统，若非如此，便无从处理数之不尽的边界情况和语言“异象”。正如谷歌翻译开发团队的成员所说，这类“基于数百万具体特征的模型，要比那些关注一般规则的精巧模型表现得更好”[59]。我们必须要对例外情况加以珍惜，绝不能随手丢弃。要知道，无论是例外情况，还是罕见情况，都包含了大量的信息。

这种机器学习技术利用的是概率和大量参数，而非原则性的规则。[60]这种尖端技术正越来越多地被应用于科学领域和其他诸多领域，从犯罪侦测到医学诊断，再到保险推销，等等。事实上，我们的审美品位也相当复杂。奈飞公司在向一个团队颁发“推荐引擎改进奖”时发现，该团队的解决方案是由各种统计技术拼凑而成的。这场比赛似乎表明，没有哪个简单的算法能够显著提高推荐的准确性；无论谁是获奖者，都需要使用更复杂的方法来捕捉和预测人们对电影的个性化需求和“怪异”品位。

这种现象其实在所有类型的技术当中皆有呈现。计算机科学家小弗雷德里克·布鲁克斯明确指出：“软件的复杂性是软件的根本属性，而不是偶然属性。”[61]

即使是法律体系这个复杂系统，也会受到例外情况和边界情况的影响。很多人都认为，在合法与非法之间有一条明确的界限，但事实并非如此。恰恰相反，边界会随着时间的推移而不断伸缩、折叠，变得凹凸不平或纠缠不清，所以边界不可能是一条明确的界限。最终，法律体系看上去会是一个分形结构：无论你将一个小的图形放大多少倍，依然会

发现很多的不均匀，依然有更多的细节需要进一步观测。[62]任何的一般性规则最终都必须应对例外情况，而后者又会裂解成更多的例外和规则，从而形成越来越复杂的分支结构。

对此，法学家杰克·巴尔金（Jack Balkin）在一篇题为《法律思想的结晶结构》（*The Crystalline Structure of Legal Thought*）的文章中阐释道：

> 我们可能会想，在关于疏忽的客观标准中，是否存在适用于儿童的例外规则；或者，是否存在适用于疯子、盲人，以及其他特殊人群的不同标准。这样的想法会引导我们开始进行接下来的规则选择，而且每个选择都会滋生出更多的法律含义分支。举例来说，假设我们遵循某个法律含义分支的发展轨迹，为儿童制定了一个例外规则，当然，现在这是一个多数规则。我们可能还会继续考虑到，当孩子从事成人活动时，是不是也需要一个例外规则，而现在，这通常也是一个多数规则。然后，我们可能会继续追问，在该规则的含义范围内，驾驶摩托车是否属于成人活动，如果是，那么驾驶小轮踏板摩托车是不是也属于成人活动？这个过程可能会一直持续下去，最终我们会得到一个关于规则选择的递减序列，但是事实上，复杂性和特异性都在不断递增……[63]

法学教授戴维·波斯特（David Post）和生物学家迈克尔·艾森（Michael Eisen）也携手研究了这个问题。[64]虽然他们无法证明，任何一个法律陈述都能被进一步细分，毕竟这是一个“乌龟背上的世界”式的命题(6)，但是他们指出：“我们确实从来没有遇到过无法分解为子问题的法律问题。”波斯特和艾森还通过模型证明，某些类型的法律分支结构实际上具有分形结构的特性。在现实中，他们对诉讼案件中的法官意见进行了分析，并在引用的法律条文中也发现了指向分形结构的特征，从而验证了模型的结论。[65]由此看来，法律体系的分形复杂性很可能不只是一个令人回味的隐喻。

正如法律学者马克·弗勒德（Mark Flood）和奥利弗·古迪纳夫（Oliver Goodenough）所指出的那样：“好合同和好律师的价值，在很大程度上基于这样一种看似烦琐的规划：对于随时可能因出轨而覆灭的婚姻关系，要如何对一切可能的出轨方式未雨绸缪。”[66]换言之，法律体系的复杂性通常源于例外情况及其引发的“并发症”。

我们所观察的技术系统，无论是法律、软件、设备，还是科学模

型，都会在例外情况和边界情况的驱动下，在吸积和交互的双重助力下，变得越来越复杂，越来越混乱。

越来越多的复杂系统

尽管我在上文中将那些导致系统越来越复杂的驱动力描述为不可抵挡的力量，但是我们还不至于束手无策。实际上，我们也进行了一些抗争。重要的是，在奋力抗争的过程中，我们看出了它们的“真正实力”，并洞晓了根除它们的难度。

以下策略似乎可以让技术系统变得有序且有逻辑。针对不同的系统，如果我们能够采用不同的方式去构建、设计、修改和重建的话，那么，对技术的控制或许真的可以实现。例如，我们可以尝试解耦某些系统，将它们拆解为更小的单位，以保证它们的相对简单性和可管理性。拥有物理学背景的社会学家、微软研究院首席研究员邓肯·沃茨（Duncan Watts）指出，在金融领域，解决由复杂性导致的崩溃故障的方法之一，是直接消除耦合复杂性。[67]这就好比，如果某家公司的规模已经变得太过庞大，且预期其失败会造成级联式冲击，那么就必须对它进行拆分，或缩小其规模。

其他学者也讨论过类似的问题：怎样才能让一个大型系统的互操作性达到最佳水平。[68]这里所说的最佳水平是指，系统既能运行良好，又不会因高度的不可预测性而产生负面效应；也就是说，我们的目标应该是创造最佳水平的互操作性，而不是最大限度的互操作性。当然，这个目标不是那么容易就能实现的。想到是一回事，做到又是另一回事。实现这个目标的方法之一是坚持使用特定的设计原则，将可理解性和模块化内置于我们的设计中。

我在前文中已经阐述过，当一个系统具有高度的互联性时，我们便很难将其拆解，也就很难探究其内部发生的一切。但是，在大型系统中，确实可能存在这样的情况：某些部分之间的互联性远强于它们与其他部分之间的互联性。换句话说，系统中存在若干模块，而每个模块由若干“部分”紧密互联而成，并在一定程度上保持独立。我们常在生物学中见到这样的模块，它们皆包含了若干“行动一致”的“部分”。植物的线粒体和人体的心脏皆是如此。当然，这些模块通过别的身体器官和化学信号，以及其他方式，和系统其他部分保持着紧密的联系，所以我绝不建议人们进行心脏摘除手术。不过，这些模块又是相对独立的，即使不

借助系统的整体功能，也可以被理解。

在技术系统中，我们同样可以观察到模块化现象。当一个软件拥有多个独立的功能或部分时；当你交替使用不同的应用程序来完成同一项工作时(7)；当你分析《美国联邦法规》中那些相对独立的部分时，你都可以看到模块化的身影。模块化遵循了抽象的原则，实质是通过将系统拆解为若干部分，从而在一定程度上实现对复杂性的管理。

然而，从单个模块入手去理解系统，或者在单个模块的基础上构建系统，也并非总能让我们如愿以偿。如果每个模块都是多向输入和多向输出的话，那么当它们关联在一起时，系统的行为很可能仍是令人难以理解或无法预测的。最终结果很可能是“组合爆炸”：由于潜在的、不同的相互作用太多，模块数量很快就会超出我们的处理能力。例如，假设系统中的每个模块都分别具有6个不同的输入路径和输出路径，同时这个系统拥有10个模块，那么将所有这些模块关联到一起的方法将比整个宇宙中的恒星还要多。[69]

在能够进行严格监管的领域里，例如，在金融系统中，或是在企业构架方面，找到理想的互操作性水平或加强模块化，是有可能实现的。我们可以规定，当机构达到一定规模时就必须进行拆分。然而，事与愿违，在大多数其他类型的技术系统中，各部分之间的互联通常只会继续猛增。当系统规模相对较小时，我们尚能进行模块化处理或分段构建，但是随着技术的发展，这种边界清晰的处理方式将变得越来越不可行。在社会压力和系统传统结构的共同作用下，我们不得不继续强化系统之间的相互联系，以致它们越来越难被分解。因此，尽管我们渴望简单，但现实却背道而驰。

虽然我们也可以尝试构建一些设计得更合理的系统，但这种方法只能在一段时期内行之有效。例如，现在业已出现的“计算机科学和工程实践”[70]，其实质是“工程卫生学”，可以大大减弱系统的复杂性，比如避免计算机程序出现某些类型的变量。如果丰田公司采取了这种做法，那么其汽车系统的整体复杂性就会减弱很多。此外，专业软件通常也会包含一些能够降低代码错误率的设定。这些方法能够将每千行代码的错误率降低至6%，而这显然是一个极低的数字。[71]当人们共同构建、操作和维护复杂的技术系统时，管理团队的一些特殊操作将有助于减少系统问题。但是从长期来看，吸积、交互和边界情况最终会使这些简化工作成为徒劳。

随着时间的推移，系统渐渐变得复杂；而面对这些复杂系统，我们

的大脑也渐渐无能为力。无论是互联网，还是大型基础设施，要想从整体上理解它们，已经不可能了。

那么，为什么必然会产生这样的结果呢？在接下来的一章中，我们将讨论人类理解能力的社会极限和生物极限：无论有多么努力，我们的大脑和社会在面对这些复杂系统时的表现都不会太好。

OVERCOM PLICATED

Technology at the limits of comprehension

对于大多数人来说，记住一个7位数就相当不容易了。当复杂系统的组成部分和连接数量猛增时，即便是专家也会望而生畏。我们个人的知识储备，与理解复杂系统所需要的知识相比，存在着根本性的冲突。

03 为什么复杂系统越来越难以理解了

对于大多数人来说，记住一个7位数就相当不容易了。当复杂系统的组成部分和连接数量猛增时，即便是专家也会望而生畏。我们个人的知识储备，与理解复杂系统所需要的知识相比，存在着根本性的冲突。

1985年的一天，一位患者来到一家诊所接受宫颈癌的放射治疗，所用的设备是Therac-25型大型放射治疗仪。[1]操作员按照医生的处方设定好剂量并开始治疗，而后放射治疗仪提示“信息错误”并称“无辐射剂量”。操作员重新尝试了一次，结果还是一样。此后，操作员又尝试了三次，也就是说，他为患者进行了五次治疗，而每一次治疗仪都提示“无辐射剂量”。治疗结束后，患者抱怨髋部有灼烧感，随后被送入医院接受治疗。

几个月后，该患者因癌症去世。后来发现，她遭受了可怕的过量辐射，即使能够幸存，她的臀部也需要被全部切除。然而，那台机器却说“无辐射剂量”。

这并不是特例，这种型号的放射治疗仪还出过其他故障。20世纪80年代，有6位患者经历了Therac-25型放射治疗仪的错误治疗，在一次治疗过程中遭受了过量的辐射。辐射严重过量对患者身体造成了极大的伤害，甚至导致其中一些患者失去了宝贵的生命。辐射过量事故是这种型

号的放射治疗仪问世以来最糟糕的故障。

能不能避免这种故障？或者降到最低？仔细看一下制造商于1983年提供的关于这些仪器的安全分析报告便不难发现，导致故障的原因之一是：参与设计和测试的技术人员只关注了硬件问题，而忽视了软件问题，因为他们认为“软件不会因磨损、疲劳或复制而退化”[2]。虽然这句话大体属实，但他们却完全忽略了一个更重要的事实：软件本身就很复杂，出错的形式也多种多样。这份报告表明，制造商对软件的复杂性缺乏认知，而这种复杂性可能会导致致命的辐射过量。软件缺陷在生活中并不鲜见，但是该制造商的安全分析几乎完全忽略了它们隐含的风险。

显然，因为技术人员对软件的复杂性有所误解，所以Therac-25型放射治疗仪的安全性未能得到保证，换句话说，是那些技术人员酿成了这场致命的灾难。事后看来，他们理应很容易找出问题所在，然而他们低估了系统中各个组成部分的重要性，最终导致了致命的故障。不过，在新的技术系统中，对问题的诊断的确已经变得越来越困难了。无论我们多么努力地构建和设计符合逻辑的技术系统，系统中总会有某些部分是我们无法完全理解的。原因很简单：我们是人。人类的思维方式与复杂系统的运行方式是完全不匹配的。复杂系统的构建方式决定了它们很难被理解，甚至不可能被理解。

在学习编程时，最初要学的技能之一，是用与日常生活中不同的方式进行计算。当然，这并不意味着必须用二进制或十六进制计数，对于大多数程序员来说，这只是一个好玩、但不必需的技能。我的意思是，程序员总是需要从“0”开始计数，在他们的列表中，第一个对象始终是“0”。之所以要从“0”而不是从“1”开始计数，是因为那就是计算机计数的方式。[3]正如作家斯科特·罗森堡（Scott Rosenberg）所说，计算机计数方式与日常计数方式之间的差别，正是我们需要对计算机代码进行调整的地方之一，同时也是错误和故障的发源地之一。[4]我们必须在计数时调整1个数位，反复增加或减少，以协调人类世界的“编号”与计算机世界的变量枚举之间的差异。如果调整失败，错误就会倍增。

人类世界以“1”为始，而计算机则从“0”开始计数。这个事实意味着：人类的思维方式与大型系统的构建及运行方式之间的裂痕，将会越来越大。我们无法一一跟踪大型系统中的所有过程，以及每一次交互所导致的一系列后果。人类的大脑并不具备解析这种复杂性的能力。在大型系统的构建方式和人类的思维方式之间，我们看到了诸多复杂因素，而这些复杂因素不仅会降低系统的可理解性，还会导致出人意料的后果

和问题。

例如，在军事行动中，战士们经常要面对非常复杂的情况。这便要求他们既要具备处理大量信息的能力，又要具备快速的反应能力。但是有些时候，如果情况太复杂、太紧张、太混乱，战士们可能会不堪重负，继而丧失处理紧急事件的能力。据说，当这种情况发生时，战士们会变得“像无头苍蝇一样”（lose the bubble）。正如加拿大政治学家、生态学家托马斯·霍默 - 狄克逊所描述的那样：“以往不难理解和可以预见的一切，突然之间都变得混沌不堪、令人困惑。”[5]由于对形势的认知急剧下降，战士们已无法处理接二连三到来的刺激信号，更无法采取行动。

这个问题不只会出现在那些身处紧张局势、面对复杂情况的人身上，例如航空交管人员，以及需要在战斗中做出冷静判断的人，其实这个问题会出现在所有人身上。在面对技术系统时，无论是个人，还是集体，都会遇到这样的问题。人类已经失去了控制权。

在纠缠世界中，我们是因为以下两种相互关联的原因失去控制权的：第一种，无法完全把握庞大而复杂的系统结构和动态，即无法把握系统的不同部分以整体的形式相互作用；第二种，无法获取足够多的必需的、用以理解这些系统运行方式的专业知识。接下来，我们将再一次以语言学为例，力求更清晰地说明，人类的思维方式为什么无法适用于复杂的技术系统。

力不从心的大脑

在计算机科学中，“递归”（recursion）这个术语本意为自我引用，它描述了一部分需要引用回自身的计算机代码。这个概念催生了不少程式化幽默。在搜索引擎上查找“递归”这个术语时，它可能会反问你：“你的意思是递归吗？”这对许多人来说无疑是很有趣的。

递归也潜藏在语言结构中。换句话说，语言具有递归能力。事实上，语言在理论上是无限递归的。你可以说“他说狗是棕色的”，也可以说“她认为他说狗是棕色的”。你在第二句话中嵌入了第一句话。更大胆一些，你甚至可以说“我记得她认为他说狗是棕色的”。基于不同的句型结构，在一个句子中嵌入另一个句子的现象，可以出现在句首、句中或句末，而且还可以一遍又一遍地嵌入。

递归，使语言变得无限丰富。举个例子，假设有一个规模相对较小的语言系统，它仅包含1 000个动词、10 000个名词，以及这样一个规则：将单词组成句子的唯一方式是“名词+动词+名词”。这种语言看似简单，但应用功能却极其强大。你可以造出10 000×1 000×10 000个句子，也就是1 000亿个句子。就算你以每10秒钟说一句话的速度，时刻不停地说，也需要3万多年才能说完所有句子。[6]更重要的是，上述事例中的语言系统，其实是一种体量很小的语言系统。如果你面对的是一种单词量更大、句型结构更复杂的语言，想要将所有句子都说一遍，所需要的时间会十分漫长，漫长到需要地质时间来描述。

虽然句子可能会多到不可思议，但其数量终归还是有限的。在上文中，我们讨论了大得惊人的有限性，现在我们需要将讨论转向真正的无限性，不过在此之前，必须再谈一谈递归这个概念。一旦你在一种语言系统中引入了递归，也就是允许将任意数量的子句嵌入到主句中，那么这种语言的丰富程度在理论上已具有了无限性。

然而在现实中，这种说法并不完全正确。说一种语言系统允许任意数量的子句嵌入主句，无疑是愚蠢的。尽管就语法规则而言这是可行的，但是实际上，我们的大脑根本无法解析那么多层的递归。尽管我们希望能拥有无限性和多样化的语言，但我们能处理的递归深度通常不超过三层。

下面这两个例句引自认知心理学家、语言学家史蒂芬·平克（Steven Pinker）(8)的《语言本能》。[7]句子不仅难以理解，而且语法似乎也是错的。

> The dog the stick the fire burned beat bit the cat.
>
> 一只被火烧过的棍子打过的狗咬了这只猫。
>
> The rapidity that the motion that the wing that the hummingbird has has has is remarkable.
>
> 蜂鸟所拥有的翅膀的运动的速度有了显著提高。

这两个句子所嵌入的层级其实都不算多。例如，第一句话的逻辑是一只狗被火烧过的棍子击打，然后咬了一只猫。这句话的语法结构是：对“狗咬猫”中的“狗”进行修饰，同时对棒子进行描述。所以我们说，这句话只有两层嵌套。如果句子的嵌套达到了10层，就不可能有实际意义了；如果人类连10层嵌套都处理不了，就更不可能处理无限嵌套了。

有一些类型的语言处理任务是人类可以完成的，比如翻译，而且人类可以做得比计算机更好，但是在解析句子方面，计算机则有很大优势。人类的认知会受到大脑工作记忆的限制，但是计算机却可以使用大容量的存储器，将句子的各个部分放在相应的位置上，然后构造出句子树，并使其拥有某种意义。正因如此，计算机才可以轻松地分析那些令人类大脑“崩溃”的句子。例如下面这种奇怪的句子：“This is the cheese that the rat that the cat that the dog chased bit ate.”（这是被狗追赶的猫咬了的老鼠吃的奶酪）。无论是谁，听到这样的话肯定都会头昏脑涨、不明所以，然而计算机却可以轻而易举地做出解析。[8]

有些计算机程序甚至直接内置了语法结构，因而能够创建出层级相当庞杂的句子。不过，它们无法让这种复杂的句子变得合理，只能让它们看上去像是人写的，并且风格与某些作家相似。以“康德发生器”（Kant Generator）[9]为例，它可以造出这样的句子：“由于对现象的知识是先验的，因此，如我所知，对读者来说，就必然性和事物本身而言，应该非常小心地观察，人类理性的规律，可以像我们的判断一样对待。”人类作为生物，比小小的计算机程序复杂得多，但是在拆解问题，并确定其合理性方面，却无法像计算机程序那样信手拈来。

与此类似，还有一种被称为“花园小径句”（garden path sentence）的语法结构。[10]例如“The complex houses married and single soldiers and their families.”这类句子，看上去很有趣，但实际上却充满歧义。这些有趣的句子一开始是单向的，但最后呈现出来的语法结构和意义却完全超出了我们的预料。当句子切换路径时，我们会自动加入既定的含义，但结果往往令人惊讶，并让人一时间摸不着头脑。

当然，除了在面对递归以及其他某些语法技巧时，我们会显得无力之外，人类认知能力的局限性还体现在其他诸多方面。事实证明，许多机器能够轻松完成的任务，人类却无法胜任，这种情况几乎成就了整个家庭手工业。因此，我们无须再讨论“视觉幻象”这类感性挑战了，不如来看看有关心理能力极限的实例。[11]例如，你有没有想过，你能瞬间记住的数字有多长？对于大多数人来说，只能记住7个数字左右，也就是说，你很难记住比电话号码更长的数字，而且这里所说的电话号码是不包含区号的。[12]

又如，你了解自己的计数能力吗？更直接地问，你能数清楚计算机屏幕上有多少个点吗？当然，人类绝对有能力数很多数，但是你一眼就能看清，并能准确地说出来的屏幕上的点又有多少呢？事实证明，这个

数字非常小。当你观察一大组点时，大脑会将它们分成多个小组，通常是三四个组。在视觉上，大多数人能瞬间识别的物体只有4个左右。这种对数量的即刻察觉能力被称为感知（subitizing）。这是人类大脑的怪癖之一：只能对少数几个对象进行有效的感知。例如，当我们在阅读冗长而曲折的句子时，通常会遇到障碍，因为我们无法同时识别4个以上的对象。

在与计算机的其他“较量”中，人类的“成绩”也很可悲。将一个信息小片段传输到大脑的长期记忆区大约需要8秒钟，但是在同样的时间，甚至更短的时间里，我们可以将整本《战争与和平》下载到笔记本电脑中。[13]人类在处理多重任务时的效率极差，远远比不上最低配置的计算机。我们的神经元回路比计算机电路要慢100多万倍。[14]另外，据估计，大脑的长期记忆容量甚至还比不上20世纪80年代初出厂的一台老式苹果机（Macintoshes）。

人类的大脑最终能不能克服其局限性呢？通过对人类认知的研究，我们得到的答案并不如意。就像可以加快计算机转速，让计算机超频运行一样，我们偶尔也可以利用药物来“提神醒脑”。但是研究表明，大脑在“超频”时也会权衡利弊，做出取舍。正如超频运行的计算机可能会发热一样，超极限运转的大脑也会受到伤害。从现有的证据来看，人类的大脑在进化过程中已经微妙地实现了优化，[15]所以随意“超频”可能会致使它出现严重的问题。

只要仔细观察一下那些拥有超强记忆力的“奇人”便可知，这种情况实则是“矫枉过正”。这些人可以记住遇到的每一件事，但是他们并不是超人。事实上，他们会遇到许多困扰，比如，在诸如“面孔识别”之类的简单任务中，他们的表现会非常差。这是因为他们的记忆过于细节化，所以只要人的面孔出现一点点变化，他们就很难进行确认。在阿根廷作家豪尔赫·路易斯·博尔赫斯（Jorge Luis Borges）的短篇小说《博闻强记的富内斯》（*Funes the Memorious*）[16]中，主人公富内斯拥有完美的记忆能力，同时也因此而背负着无比沉重的负担，因为每一个细节变化都会在他的大脑中形成新的记忆。就像富内斯这个虚构的人物所面临的处境一样，在现实生活中，这种令人难以置信的完美记忆能力会导致其他诸多方面的问题，例如缺乏抽象思维能力。若一个人不得不背负大量的不必要信息，一切就会变得得不偿失。

在这个世界上，的确存在拥有超常认知能力的人。例如，有的人拥有超大记忆容量，有的人能够心算超多位数的加减乘除。既然如此，那

么在理解复杂系统这件事上，应该也能找到极具洞察力的人。在这一点上，印度数学家斯里尼瓦瑟·拉马努金（Srinivasa　Ramanujan）就是最典型的例子。[17]拉马努金是一位自学成才的数学天才，可惜英年早逝，如同一颗彗星般匆匆划过了20世纪早期的天际。他不是一位普通的数学家。普通的数学家只能依靠反复试错和偶尔闪现的灵感来解决问题，但是在拉马努金那里，方程式会像瀑布一样直接从大脑中奔涌而出，而且除了少数方程式不完全正确以外，绝大多数方程式都复杂和准确得令人震惊。

苹果公司创始人之一斯蒂夫·沃兹尼亚克（Steve　Wozniak）堪称技术研发领域的拉马努金。他不仅完成了对第一台苹果电脑的编程，还负责了第二代苹果电脑的开发。正如程序员兼小说家维克拉姆·钱德拉（Vikram　Chandra）所说："这台计算机的每一个组成部分、每一个字节都是沃兹尼亚克一手打造的，而且人们尚未发现任何错误……沃兹尼亚克一人包办了所有的硬件和软件，还在计算机代码中创建了一种编程语言。他本人就是硬核。"[18]言下之意，沃兹尼亚克对技术的理解无人能及。

事实上，在现实生活中，我们也能观察到大脑的认知极限以及这种极限的扩展方式。伦敦的出租车驾驶员对道路知识的获取和运用，便是一个很好的例子。"知识"是一个奇妙的术语，这里所说的知识是指伦敦地区道路系统的所有细节：25　000条道路及其相互之间的连接点，以及公共汽车站、地标性建筑、雕像、餐馆、酒店……[19]为了将乘客准确高效地送达目的地，出租车驾驶员必须掌握全部细节。积累这方面的知识，成为一名出色的出租车驾驶员，可能需要长达数年的强化记忆和积极的实践。最终，这些出租车驾驶员的大脑出现了明显的变化：对空间记忆至关重要的海马后部脑区增大了。

然而，尽管存在这样一些"异常点"，尽管这些特例给我们留下了非常深刻的印象，但是人类大脑的认知能力终归还是有限度的。[20]拉马努金也会犯错，沃兹尼亚克肯定也有无法突破的认知极限，譬如递归问题。而伦敦的那些出租车驾驶员，显然也不可能把全世界的道路信息都存在脑子里。

如前所述，我们的记忆能力以及在记忆中进行检索的能力，都是有限的，能瞬间识别的事物数量也少之又少。除此之外，我们还很难理解系统内部各种互联的具体含义。具体来说，在面对非线性变化时，我们会不知所措。当某个事物以线性方式发生变化，即一个较小的变化产生

一个较小的差异、一个较大的变化产生一个较大的差异时，我们所需要做的主要是以线性形式进行外推。在这样做的时候，大脑几乎不会遇到什么困难，因为线性系统的输入与输出是成正比的。但是，在非线性变化中，一个小变化在流经一个大型互联系统时，会导致系统以不成比例的方式发生变化，以致大脑无法很好地做出应对。

在非线性系统中，行为会受到反馈和输入放大率调制的影响（或者也可能相反：一个很大的变化只能带来一个极小的影响），这令大脑很难将输入与输出关联起来。在这种情况下，我们无法继续以线性形式进行外推。所有变量之间的互动形成了错综复杂的不规则曲线，让我们的大脑一筹莫展。正是因为大脑本身具有这样的缺陷，所以我们很难把握复杂系统，包括我们自己所构建的复杂系统。

认知的极限

哲学是一门泛科学，无论是政治还是伦理，抑或是科技，都属于哲学研究的范畴。在有关科技的哲学研究中，人们对“软件的哲学含义”这个命题越来越感兴趣：我们应该怎样理解软件这种算法作品？为此，美国堪萨斯大学的哲学家约翰·西蒙斯（John Symons）和杰克·霍纳（Jack Horner）研究了这样一个问题：为何我们对软件这种技术系统的构建，几乎立即就会产生无法理解之感。[21]

这种情况之所以会出现，最直接的原因在于：软件中存在大量的分叉点。所谓分叉点指的是：如果条件A为真，那么就做这一件事情；如果条件B为真，那么就做那一件事情。对程序员来说，这叫作if—then语句。例如，有的计算机程序可能会规定，如果某个数是奇数，那么就加上10；如果是偶数，则加上5。

正如西蒙斯和霍纳所指出的那样，一旦某个计算机程序包含了多个分叉点，那么该程序运行路径的数量会倍增。据保守计算，一个只有1 000行代码的程序（即使是对于一个非常简单的程序来说，1 000行也是一个很少的数量；实际上绝大多数程序的代码行数都不会小于这个数字），当分叉点以正常频率出现时，也会有10^{30}个潜在路径可供遍历。10^{30}是一个天文数字，超过了1亿亿亿，因此要检查所有可能的路径，理解每个路径的含义，并确定其准确性，不仅是不可行的，也是不可能的。可见，这个系统的确是很难被理解的，即使从宇宙诞生之初便开始研究，也无法掌握它的所有细节。换句话说，绝大多数计算机程序都是

永远不可能被任何人完全理解的。这个结论适用于笔记本电脑中的应用程序、厨房设备中的计算机代码，以及用来调度全球飞机的航空软件……

当然，在某种程度上，计算机程序也具有一定的可理解性，而且不需要我们手动遍历所有潜在路径。这正是抽象的主要特点之一（请参见本书第1章的讨论）。在抽象的基础上，对软件进行各种严格的测试和纠错，并遵循软件“卫生准则”，比如，禁止使用GOTO语句或某种类型的变量，就可以有效缓解理解无力的状况。但是我们仍然无法确定，我们是否真的了解所有的影响和结果。无论是使用某个科学模型的科学家，还是操作大型机器的技术人员，或是驾驶最先进汽车的驾驶员，都不得不适应这种“不完全理解”的状态。这是生活在纠缠世界中必须付出的代价。

通过大数据分析，我们可以发现庞大和互联所带来的影响。大数据分析就是用特定算法对海量数据进行处理，以增强预测能力，但有时会以牺牲可理解性为代价。[22]例如，谷歌公司已经开始借助算法来提高效率了，[23]其中就包括通过将诸如冷凝器数量、水泵数量以及室外风速等各种数据信息输入特定的计算机模型，来提高其数据中心的能效。在谷歌公司的博客上有这样一段话：“在像数据中心这样的动态环境中，人类很难观察到所有变量……的相互作用，而计算机则特别擅长发掘隐藏在数据中的潜在故事。为此，数据中心的工程师们将日常运营过程中收集到的信息输入了模型，让模型来帮助他们理解作为凡人可能无法注意到的复杂交互。”[24]

要搞清楚这种大规模技术模型的运算细节是非常困难的。而且，正如《新科学家》杂志（*New Scientist*）首席技术编辑道格拉斯·希文（Douglas Heaven）所说，即使我们能够做到这一点，也不一定有意义。[25]这种软件系统所给出的选项或答案，并不是按照人类的思维方式得出的，而且常常有违于某些稳定的一般规则或思想。这些决策所基于的路径并非简单的逻辑推理，而是一系列异常复杂的运算。我们置入大量的信息和数据，让庞大的软件系统“生产”出一些东西来。然后，可以得到了一个答案，它可能确实有用，但同时我们也失去了对解答过程的洞察力。从法律体系到硬件设施，在一个又一个的领域中，我们始终在通过与计算机“合作”来驾驭复杂无比的技术。然而，在这个过程中，我们又对自己所依赖的这些系统的运行方式感到困惑不已。

不仅如此，这个过程还在一直加速。在进化计算领域，软件已经可

以自行“进化”出问题的解决方案了，但我们对最终方案的形式几乎一无所知。那么，当我们需要一个公式来拟合数据的时候，公式又从何而来呢？我们需要借鉴生物学中的进化思想。先在计算机程序中创建一个由潜在的解决方案组成的“种群”，然后让它们自行进化，也就是让它们重新组合、变异和复制，直到最适合的解决方案脱颖而出。进化算法可以很好地实现这一过程，即使你无法理解它给出的最终解决方案，也没什么关系。

这方面的研究多年以前就已出现，其中一项研究是利用进化算法，设计特定类型的计算机电路。[26]具体来说，先选定特定任务所需的电路类型，然后尝试在硬件中进化出一个解决方案，也就是让一组备选电路混合在一起，形成一个达尔文式的“进化浓汤”。经过多重进化之后，计算机程序就会发现一个最完美的电路设计。这种电路设计有一个奇异的特征：有一部分设计是与主电路断开的；而这对整个电路的功能来说至关重要。这个由进化而来电路设计利用了一些奇怪的物理知识和电磁现象，而这些知识和现象是从未有人想到过和利用过的。

此外，人们还利用一个进化来的公式解决了另一个难题。这个难题一度被认定为是不可能完成的任务。凯文·凯利（Kevin Kelly）在其著作《失控》（*Out of Control*）中是这样描述这个公式的：“它不仅丑陋，而且杂乱，甚至不可理喻。即使是对数学家或计算机程序员来说，这个进化而来的公式也是不折不扣的‘荆棘路上的柏油娃娃’。”[27]最终，研究者们还是参透了进化生成的代码，但是发现它对问题的解决方式“绝非人类所为”。这种进化生成了新颖的技术系统，但是我们很难理解这种系统，因为我们决然想象不出同样的东西。这种系统与我们擅长构建的那些系统有着本质区别。

在物流领域也出现了一些强大的算法。[28]因为进化生成的运货线路看上去不合逻辑，所以驾驶员们对这些违背直觉的线路很不满意，但最终又不得不臣服于它们的效率。事实早就证明，在国际象棋比赛中，计算机比人类强大得多：计算机能够“想到”人类棋手不可能想到的走法，并取得最终胜利。那些极具计算机特色的走法被称为“计算机走法”，人类棋手几乎从来不会使用。它们看上去“丑陋”，但效果却很好。正如经济学家泰勒·考恩（Tyler Cowen）在其《平均时代的终结》（*Average Is Over*）一书中所提到的那样，这种走法乍看上去似乎是错的，但其实非常有效。[29]一位参与IBM公司深蓝（Deep Blue）计算机设计的工程师也曾说过，“深蓝”在与加里·卡斯帕罗夫（Garry Kasparov）进行国际象棋

比赛时曾用了一着超级怪异的走法，以致“评论员们有的赞不绝口，有的大加抨击”。[30]但是人们后来发现，“深蓝”这一着既怪异又极其出色的走法，其实是程序中的一个缺陷所致。无论如何，这样的人机对弈揭示了一个事实：在最高级别的对弈中，对人类来说，国际象棋还是太复杂了，因为交互的部分太多，即使是像卡斯帕罗夫这样的国际象棋特级大师也无法完全理解。许多时候，我们甚至无从分辨哪个决定是错的，哪个决定是对的。

那么，法律系统的情况又如何呢？在法律系统中，我们也看到了同样的问题，而且这些问题正愈演愈烈。在上一章中，我们曾经引用过马克·弗勒德和奥利弗·古迪纳夫的观点。除那之外，他们还指出，“但凡对法律进行解释，那么就算解释得再明确、再精准，人们对法律文件的理解过程也会变得缓慢而曲折，[31]因为我们需要不断地核查、重新定义、交叉引用、处理例外……，如此一来也就嵌入了复杂性。但是，律师的智慧——作为传统的契约解释的计算机制，不但非常昂贵，而且会受到认知水平的限制。”因此，当我们发现如下现象时不应感到惊讶：假设我们将同一份家庭收入数据分别提交给45名税务人员，他们会反馈45种不同的纳税方案。[32]

除了系统组成部分及其交互的复杂性超出了我们的处理能力，并会导致我们无所适从之外，摆在我们面前的还有另一个不利因素，那就是我们可以保存在头脑中的知识总量也是有限的。这不仅涉及原始数据，还涉及专业技术知识。技术系统所涉及的知识面变得越来越广泛，而且还跨越了不同的领域，要知道，即使是专家也不可能了解所有这些知识。

为了了解认知的局限性，我们需要先了解一下人类追求专业化的历程。然后我们就会明白，当今这个“不可理解”的新时代，其实由来已久。实际上，人类社会长期以来一直保持着这种进程，只不过现在才抵达终点。

最后一个无所不知的人

在我的书架上有三本书[33]的书名或副书名都用到了“最后一个无所不知的人”（The Last Man Who Knew Everything）这个短语。第一本书是讲述阿塔纳斯·珂雪（Athanasius Kircher）的生平和学术贡献的文集。珂雪是一位生活在17世纪德国的耶稣会牧师。今天的人们普遍把珂

雪视为一位怪杰，但是也有一些人认为他是一个骗子。珂雪的著作可谓包罗万象，从天文学到埃及象形文字，不一而足，甚至还包括了以猫为题材的音乐剧。第二本书和托马斯·扬（Thomas Young）有关，扬出生于1773年，研究的领域涉及物理学、医学和语言学，等等。第三本书则是约瑟夫·莱迪（Joseph Leidy）的传记，他生于1823年，住在美国费城，是一位古生物学家和博物学家。

那么，最后一个无所不知的人到底是谁呢？我不知道。事实上，应该没有人能洞晓人类文明所生产的一切。在过去的几个世纪里，知识体量呈现出爆炸式增长，即使只是想要了解所有新知识也是不可能的。同时，在我们眼中，宇宙也变得更加复杂难懂了。不过，在“最后一个无所不知的人”所生活的时代，甚至在更早之前，确实存在不少努力让自己变成无所不知的人。这种努力往往和一种被称为“珍奇柜”（cabinet of curiosities）或“珍奇博物馆”的事物密不可分。

“珍奇博物馆”一词，德语为“wunderkammer”，指的是一种特殊的房子，其主人为了掌握世界上的“全部知识”而四处搜集珍奇事物，最后把房子塞得满满当当。珍奇博物馆可谓是怪异物品的庞杂集合，动物标本、草药、绘画，应有尽有。通常，它们的主人是富有的欧洲贵族。珍奇博物馆既是社会地位的象征，也是了解宇宙及其间奇迹的窗口。那些收藏家对人造事物，例如乐器和武器；以及天然物品，例如化石和矿物，兼收并蓄。他们的橱柜里装着一整个世界，而世界的多样性通过这些摆在眼前的藏品得以集中呈现。正如作家菲利普·鲍尔在其《好奇心》（*Curiosity*）一书中所描述的那样：“理想的收藏必定是全面的。这种全面性不仅是因为它包含了世界上每一种物体和物质的样本，或者至少是在向穷尽所有事物这个目标靠拢，还因为它是一个完整的世界缩影，也就是说，它以微缩形式表现了整个世界。”[34]

在珍奇博物馆中，你只要看上一眼，就可以窥见整个宇宙及其间的一切现象。但是人们很快发现，有些珍奇博物馆似乎只是大杂烩般的存在；更加重要的是它们永远不够大。菲利普·鲍尔还在书中引用了法国作家帕特里克·莫里斯（Patrick Mauries）的观点，在美洲大发现之后，人们意识到，任何一个博物馆都不再可能包罗世界万象了，譬如不可能包含所有的物种。[35]自那之后，人们不得不承认，这个世界实在太多样、太复杂了。于是，人们开始做起选择题：哪些东西应该收藏进博物馆？哪些东西可以忽略？

此后，珍奇博物馆仍然存在了相当长的时间。几十年前，我参观过

尼亚加拉大瀑布博物馆。它是当时世界上仅存的几家珍奇博物馆之一，其所有者是我一位朋友的父亲。当然，它现在已经闭门谢客了。在“大自然的怪物”展厅中，摆满了千奇百怪的动物标本，有五条腿的奶牛，还有两个头的羊。凝视着那些突变体，以及镶嵌在墙上的昆虫标本和埃及木乃伊，我深刻地体会到了“世界之大，无奇不有”这句话的含义。

但是，这些珍奇博物馆所陈列的事物远非世界之全部，想要包罗万象是不可能的。每一家珍奇博物馆的存在都意味着某个特定的选择过程已经结束。而且，在这种选择中，我们可以看到专业化分工的深化痕迹。随着知识的范畴逐渐超越了地域、文化和心灵的界限，想要掌控好周围的系统，我们就必须求助于专业化。换句话说，我们必须更好地理解细分领域，比如先进武器，或者某个学科的子领域。当然，这种转变绝非一日之功。

在几个世纪以前的很长一段时间里，有许多人仍在试图理解周遭世界，而这种努力不再只是简单的收集过程。事实上，他们都博学多识，例如生活在17世纪后半叶至18世纪前半叶的哲学家、科学家和数学家戈特弗里德·莱布尼茨。根据史学家丹尼尔·布尔斯廷（Daniel Boorstin）的说法：“在26岁之前，莱布尼茨为神圣罗马帝国制订了一项法律改革计划，还设计了一台计算器，并试图游说路易十四不要攻击莱茵兰（Rhineland），而去修建苏伊士运河。”[36]用腓特烈大帝的话来说，莱布尼茨一个人就是“一所完整的大学”。[37]与此类似，牛顿通过万有引力定律将物体的下落轨迹、火星的运行轨道等所有现象都统一了起来。

同一时期，英格兰最古老的学院格雷沙姆学院（Gresham College）则致力于举办各种公开的主题讲座，并在天文、几何和音乐等学科领域均拥有小型的教师团队。但事实上，格雷沙姆学院的各种头衔并没有太大的实际意义，一些教授会根据可以得到的房子的质量，而非自己的专业特长来选择头衔。[38]在那个时期，专业化并不被人看重。对此，数学家伊萨克·巴罗（Issac Barrow）曾说：“一个不博学的人，不可能是一个好学者。”[39]

然而，时至今日，知识的积累和扩展已经远远超出了任何人所能掌握限度。想要在知识最前沿构建关于世界以及新技术系统的新模型，我们就必须尽可能地做到“对越来越小的领域知道得越来越多”，也就是要在某个特定领域内进行深入的专业化研究。[40]对此，美国西北大学的本杰明·琼斯（Benjamin Jones）提出了“知识负担”理论[41]：想要在前沿知识领域取得进展，你就必须先了解相关领域以往积累起来的全部知识。

由于人类的集体知识一直在不断增长，所以“知识负担”只会越来越重。在这种情况下，为了能做出新的贡献，我们不得不进行更多的学习。在另一篇与他人合作的文章中，琼斯还指出，1639年，约翰·哈佛（John Harvard）在临终前将财产和藏书遗赠给哈佛大学的前身，这也是哈佛大学名称的由来。[42]当时，他的所有藏书总共只有320本。而在今天，仅仅是美国国会图书馆就保存了超过3　600万册藏书，以及其他印刷资料。[43]“知识负担”从来没有像今天这样沉重过！当我们试图构建或理解复杂的系统时，不仅需要了解更多，而且还需要掌握日益专业化的知识。

对于这种变化，生物学家爱德华·威尔逊（E. O. Wilson）(9)是这样描述的：

> 1797年，在托巴斯·杰斐逊（Thomes　Jefferson）担任美国哲学学会会长期间，所有的美国职业科学家和人文学者都可以舒适地坐在哲学大厅的同一个讲堂里。他们中的大多数人都有能力参与讨论知识世界中的任何一个问题。当时的知识世界还很小，尚能作为一个整体来把握。而到了今天，他们的继任者——45万名拥有科学和工程学博士学位的专家，如果都来到费城，那这个城市一定会因过度拥挤而瘫痪。一般来说，专业学者只需要在自己所研究的领域内，就专业知识和研究议程与同行进行切磋；而一名成功的学者则需要在膜生物物理学（membrane　biophysics）(10)、浪漫主义诗歌、美国早期历史，以及其他类似的正式研究领域内开启自己的职业生涯。[44]

除了知识在扩展和分化，学者的数量在增加之外，各研究领域的专业化程度也在大幅提高。

我们身陷两难境地。为了更多更好地了解这个世界上的复杂系统，比如人体，我们需要将传统医学分解为众多专业化的医学学科。与此同时，我们正在构建的系统，或者说使世界运行起来的技术，却日趋庞大和复杂。于是，我们又不得不将许多不同的专业领域“缝合”到一起。举例来说，金融系统的构建需要物理学家参与；计算机系统的开发也需要经济学家参与。又例如，无人驾驶汽车的设计有赖于软件、激光、汽车工程、数字测绘等领域的专家的通力合作。[45]

换句话说，专业化帮助我们不断取得进步，同时我们也更加依赖于跨领域“汲取营养”的系统。这就要求我们必须对所有相关领域了如指

掌。然而在今天，任何人都不可能拥有全部知识。这也就意味着，无论对谁来说，这些系统在整体上都是不可理解的。

解决方案之一是推进多学科和跨学科的团队合作：将不同领域的专家组织到一起，这样就有可能在前沿地带取得突破，进而构建出特别强大的复杂系统。在软件开发领域，尽管有一些技术系统是由一个人，或一个小团队创建的，但在更多的时候，创建工作需要大规模团队的长期合作。不仅如此，在这个过程中，会不断有人加入和离开。如果将团队工作可视化，[46]也就是将关键软件的开发过程用信息图表的形式呈现出来，我们就会发现，这个过程看起来就像是一个由文字、会议和分叉组成的错综复杂的流动集束。在分叉点上，不同的个体来来去去：加入进来，参与软件开发，共同处理不同的文件，然后离开。因此，作为这种过程的产物，软件不仅非常复杂，而且往往极其庞大，以致几乎没有人能够完全理解。我们会得到这样的结果并不奇怪，因为完全了解某个特征的人可能早就离开团队了。

专业化无疑是一个成功的进程，它给我们带来了大量的、令人印象深刻的技术，但同时也将我们带入了纠缠世界。在纠缠时代里，我们不得不依赖于生而为人终无法拥有的、复杂的技术系统知识。事实上，没有人能拥有这种知识。人们千方百计地想要走出这种困境，譬如，有人说，是时候召回博学者和多面手了，应该让他们在当今时代重获新生。我们将在后文中再度审视这种可能性，在这里，我们必须先认识到：在个体所能处理的知识体量，与其需要了解的、与生活息息相关的系统知识体量之间，存在着一条无法逾越的鸿沟。

不幸的是，我们常常忽视这种不匹配；等醒悟过来，为时已晚。我们构建了大量复杂的技术系统，并确信构建的基础是符合逻辑的，直到它们迫使我们面对出人意料的结果——错误和故障，而这些结果会导致诸如全球金融市场之类的重要系统陷入混乱，甚至崩溃。有些系统会做出一些离奇的行为，这些行为甚至连设计者本人都不曾预料到，这样的系统被称为“技术狼人”。在计算机科学家小弗雷德里克·布鲁克斯眼中，大型软件项目自带难以控制的“狼人”倾向：“在我们的民间传说中，有很多令人噩梦连连的怪物，其中最恐怖的就是狼人，因为他们会出其不意地从熟人变身为恐怖的怪物。”[47]

当今时代，“狼人”就是我们自己所构建的系统突然出现的意外行为。它是所有使系统变得复杂难解的邪恶力量的集中体现。在接下来的一章中，我们就来讨论一下相关的问题。

OVERCOM PLICATED

Technology at the limits of comprehension

bug并不都是能够找到确切起因的那些错误，在以连接和交互为特征的复杂系统中，经常会出现一些令人费解的bug。尽管我们对这些bug没什么好感，但它们却是这个纠缠时代中无法回避的存在。

04 令人费解的 bug

bug并不都是能够找到确切起因的那些错误，在以连接和交互为特征的复杂系统中，经常会出现一些令人费解的bug。尽管我们对这些bug没什么好感，但它们却是这个纠缠时代中无法回避的存在。

在20世纪80年代，游戏《小蜜蜂》（*Galaga*）非常受欢迎。这是一个经典的射击游戏，游戏者可以通过操控自己的飞船来射杀所有“敌人”。作为“古老”的游戏之一，它不仅画面单一，而且配音笨拙，然而这一切丝毫没有阻碍它的流行。不仅如此，它还有一个有趣的bug。如果你在进入游戏后立刻消灭了绝大多数的“敌人”，并在接下来的大约15分钟内避开了剩下的“敌人”，那么，这些“敌人”就再也无法对你开枪了。这是一件奇怪的事情，尽管游戏者常常利用它来取得高分。[1]

为什么会出现这样的情况呢？许多人认为，这是因为在某些特殊时刻，控制“射击”的代码出了故障，“忘记”了刷新。不过也有人说，这是开发者有意设置的一个功能，目的是便于开发者进入街机并获得高分。虽然后面这种猜测不无道理，但是我认为可能性不大，因为在其他游戏中，这种所谓的隐藏功能和作弊入口会“表现”得“干净”得多。这可能就是一个bug。

这个bug，以及其他无数类似的bug表明，我们并不完全了解自己所构建的系统。要想了解清楚在《小蜜蜂》这款游戏内部究竟发生了什

么，为什么会出现问题，我们需要付出巨大的努力。尽管这款游戏的图形和规则看上去都很简单，但这并不代表它的复杂程度只是仅此而已；只有在它出现故障时，游戏者才能洞察到它的“真正实力”。在这里，bug不仅会指出需要修复的问题，还会告诉我们，我们正处于纠缠时代。[2]

早在1950年，艾伦·图灵（Alan　Turing）就已指出，机器能够并且肯定会通过它们的行为，给人类带来重重“惊喜”。[3]现在看来，这种“惊喜”发生的频率将会不断加快。随着人类所构建的系统越来越复杂，系统的实际行为和预设行为之间的分歧也会越来越大。这种超出预期的行为，正是前文所述问题的具体表现。通过上一章的讨论，我们已经清楚，人类大脑与生俱来的若干局限性进一步推动了系统向复杂难解的方向发展。这就意味着，我们所有人，包括终端用户和业内专家，在看到火箭升空后很快自爆，或是精心制定的不同法律条文产生冲突时，都会倍感惊讶。错误和故障是系统复杂性的不良产物，是超预期的、不被接受的东西。

我们在上一章中引述过哲学家约翰·西蒙斯和杰克·霍纳的观点。他们对软件的开发过程进行了深入地研究，并重点关注了软件系统中那些通常不会被检查，以及不会被详细检查的方面。例如，在进行计算时，数据如何存储，怎样四舍五入，等等。这些最容易被掩饰、被忽略的方面，往往最容易滋生大量问题。在一个被广泛使用的引力模拟器中，人们发现，许多错误都与一种特定的数字处理方式有关。具体来说，在该程序中，一段只有3万行的代码，出现了大约1万个同类错误，而将这些错误修复之后，模拟结果便大为不同。[4]

一而再再而三，我们发现，正是日益增强的复杂性导致了技术系统故障频出。当一种技术变得异常复杂之后，“狼人”就会出其不意地出现，并带来未知的影响。[5]在这些bug中，有一些错误是怪诞的，不过无伤大雅，譬如《小蜜蜂》游戏中的那个小bug；但是，也有很多错误是极具破坏性，甚至致命性的，例如“心脏出血”（Heartbleed）这种漏洞。这个错误主要出现在加密软件中，两年多以来，它一直威胁着全球2/3以上在线网站的安全，包括了Facebook和谷歌。

当然，说到bug，就不得不提到微软公司的Windows。1996年，Bug Detective出版了《Windows 95 Bug大全》（*The Windows 95 Bug Collection*）。这是一本专门讨论Windows 95的系统bug以及潜在解决方案的书。[6]这本书针对各种bug提出了一些解决方案，这些方案要比其

他方案更容易管理。例如，Windows的某个版本在启动时会显示错误提示，而解决方案是：直接忽略提示信息。还有些bug的解决方案则需要更加激烈的干预手段。例如，如果你的计算机安装了某种专用控制卡，那么某些程序就可能无法在Windows 95系统中正常运行。那么，他们建议的解决方案是什么呢？答案是："选择一台速度较慢的电脑，或者使用不同的柏努利控制卡（Bernoulli controller card）。"显然，这个问题并没有得到根本解决，或许，这个问题本就是无解的。

当系统庞大到一定程度时，上述情况就会发生。系统会以意想不到的方式与用户、其他系统，以及自身进行交互。事实上，在软件规模日益增长的同时，错误率也在大幅增加。不过，你不能就此认为，软件规模翻倍，每千行代码中的错误数量也只是会翻倍而已。[7]事实绝非如此。据估计，和拥有5 000行代码的程序相比，拥有1万行代码的程序在错误数量上是前者的4倍。

这种不可预测性和脆弱性，实际上是我们所构建的复杂系统的标志，虽然复杂系统对预料之内的冲击通常拥有令人难以置信的稳定性。这里所说的"预料之内"指的是系统拥有针对某种特定冲击的设定，但是在面对预料之外的冲击时，复杂性就会变成一种负担。

为了能更好地理解这种特殊情况，人们开发出了一个数学模型，即高度最优化容限（highly optimized tolerance）模型(11)。虽然系统在经过优化后可以适应各种各样的情况，但是任何"新异事物"都有可能让它们出现灾难性的故障，甚至崩溃。以波音777为例，[8]这种大型飞机是一台极其庞大的机器，包含了150 000多个子系统模块，所有模块指向的目标都是：确保正常飞行，并应对各种情况。但是很显然，它无法应对所有的意外情况。据业内专家称："波音777在应对大规模的气流干扰、载重和燃料的变化、边界层的湍流流动，以及材料的老化与不均匀等情况时，具有很强的稳定性，但这种稳定性可能会因少数超大规模的集成芯片的细微变化，或者某些软件故障而失效，从而导致灾难。"[9]换句话说，随着系统变得越来越复杂，再细微的刺激都有可能引发灾难。其实我们根本不知道未来可能会发生什么。

事实上，这些意想不到的后果与边界情况和例外情况有关。世界很大很复杂，所以需要一个更大更复杂的系统来管理它。很多情况虽然都具有偶发性，非常罕见，但是却极有可能导致技术故障，因为总体数量实在太多，而且无法被一一测试。还是以丰田汽车为例，我们不可能对它的软件系统进行全面且彻底的测试。正如计算机科学家菲利普·库普曼

所说："常规的车辆检测根本不可能找出所有不寻常的故障。"[10]这就好比，一个人就算穷尽一生也不可能遍历所有可能会发生的事故。

在超越大脑极限的复杂世界粉墨登场之后，噩梦随之而来。当然，这场噩梦并不是指具有自我意识的天网（skynet）已向人类宣战，而是说系统已变得越来越复杂，越来越混乱，以致各种故障接踵而来，不管人们能否预料到。复杂性注定会带来意想不到的后果，而我们却只能在问题出现时才意识到。

并不是所有bug都能被消除

哥伦比亚大学城市发展学教授凯特·阿舍尔（Kate Ascher）出版过一系列有关城市建设、交通网络和大型建筑的著作，并在书中讨论了这些系统的复杂性。[11]这些著作附有很多信息丰富的图表，细节描述也很生动有趣，但是读者读起来仍会感到些许压抑。所有这些系统都已历经了几十年，甚至几个世纪的吸积。随着时间的推移，新的部件一层层地被叠加上去，从交通网络到新能源网络，再到物流网络，皆是如此。例如，为住宅和公共场所供水和排水就是一件非常复杂的事情。为了能有一个直观的印象，让我们来看一下纽约市排水系统的巨大规模：仅在市区范围内，地下管道的总长度就超过了9 700千米，而这还只是这个精心设计的排水系统的一小部分；事实上，这个排水系统的日排放量超过37亿升。

然而，通常只有在出了问题之后，我们才会意识到系统的复杂性。2010年春，因为一场事故，波士顿大都会区的居民们接受了关于如何管理和分配水资源的"速成培训"。在那年5月的第一天，马萨诸塞州韦斯顿市的一条主供水管破裂，[12]而该水管的水是从阔宾水库（Quabbin Reservoir）输送过来的。在接下来的几天里，包括我当时居住的布鲁克林区在内的许多社区的居民都收到了通知：必须把水烧开后才能饮用，因为"现在是用备用水源在供水"。除了对面的剑桥市，水库周围的城镇无一幸免，这是因为剑桥市拥有自己的独立水源。市政工程人员当然知道排水系统的复杂性，但是对于大多数城市居民而言，只有在系统发生故障之后，才会认识到这个事实。

安德鲁·布卢姆（Andrew Blum）在其著作《管道》（*Tubes*）中对互联网的物理基础设施进行了探讨。[13]他以亲身经历作引：有一天，他家里断网了，原因是后院的网线被一只松鼠啃断了，于是他开始琢磨互

联网的物理连接特性，也就是一种纵横交错的有形网络。

在开源软件开发行业，“林纳斯定律”（Linus's Law）被许多人奉为真理。这个定律是以Linux系统的创始人林纳斯·托瓦兹（Linus Torvalds）的名字命名的，意思是“只要给予足够的关注，所有的bug都不是问题”。换句话说，如果能有足够多的人去检验某种技术，那么任何故障，无论它有多么复杂，看上去有多么难以处理，都是可以修复的，因为总会有人能找到解决办法。

但是，随着系统变得越来越复杂，这个“真理”似乎不再成立了。并不是所有的bug都可以消除：当我们面对复杂的充满交互的系统时，发现并消除每一个bug的可能性微乎其微，并且每一次修复都会引发新的问题。[14]这听起来非常令人沮丧，但在某种程度上来说确实是这样的。幸运的是，我们至少还有一丝摆脱部分困境的希望。

“技术狼人”不仅是人类跨入新时代的标志，同时也为人类指出了管理复杂系统的新方向。波士顿大都会区的供水危机说明了，自来水不是“自动来的”，它也有真实的源头。不断地检测bug是我们了解这个纠缠时代的性质，并保证自身在此间繁衍生息的有限选择之一。

从错误中学习

几年前，谷歌公司的电子邮件服务系统Gmail出现过一次严重的服务中断故障，导致许多用户在大约18分钟内无法登录邮件系统。[15]调查结果表明，问题出在谷歌软件的一个小更新包上。那是一个用来平衡邮件处理流量，以保证整个系统不会有任何一部分过载的软件包。软件包中的错误导致许多运行正常的服务器被认定为不可用。虽然这个错误并没有影响到谷歌的其他服务，但是由于Gmail需要特定的数据中心信息，所以它直接崩溃了。引发这个级联式故障的是一个很小的问题。对于这种级别的小问题，很少有人会预料到它会导致如此严重的系统故障。由此可见，系统中隐藏的某些互联性只能在发生故障时才会显现。

在调试一项技术，或者在试图根除某个错误时，你会发现，系统的实际运行方式与你所期望的大为不同，无论是在汽车软件、互联网安全程序，还是城市基础设施中。在某些情况下，出现的错误都很简单，很容易被理解和修复；但是在更多的情况下，错误是不易被觉察的，甚至几乎不可能被诊断和修复。重要的是，对这种错误进行分类和编目，恰是我们研究复杂系统某个部分的第一步。这部技术世界的“博物志”至关

重要，就像博物学家走进大自然、研究大自然，将物种及其复杂性分类编录一样，我们也需要对技术采用类似的研究方法。

我们将会越来越需要“技术博物学”。在这里，不妨再来看一个有关编程的例子。假设，你心中想着某个介于1至100之间的数字，而我的任务是尽可能快速并准确地猜到它。我会先问你，这个数字是不是超过了50；如果是，那么我会接着问你，它是大于75还是小于75……以此类推，我不断地将剩余数字一分为二，直到猜出你心中所想的那个数字。这种方法被称为“二分检索法”（binary search），其核心是将待搜索的数字分为两组。众所周知，这是一种在大型的有序列表中查找所需内容的高效方法。

在整个软件世界中，二分检索法的不同实现形式随处可见。因此，我在2006年读到谷歌公司所发布的一篇博客文章时深感不安。[16]这篇文章讨论了二分检索法的多种实现形式，并指出了它们共同的失败之处：“号外！号外！请务必认真阅读本文内容：几乎所有的二分检索法……都崩溃了！”

虽然在普通软件中执行二分检索法的代码通常都能正常运行，但是事实证明，在涉及大量数据的情况下，二分检索法的许多实现形式都有可能会失败。那篇博客文章这样写道：“当数组的长度，也就是数组元素的个数达到了2^{30}及以上时，或者说多于10亿个元素时，二分检索法就会出问题。如此海量的数据在流行《编程珠玑》（*Programming Pearls*）的20世纪80年代，是不可思议的事。《编程珠玑》是一本介绍计算机编程方法的经典著作，它所提出的二分检索法的实现形式也包含了这个错误。但是到了今天，在谷歌公司和其他地方，海量数据都已经很常见了。”

直到今天，在庞大的数据集随处可见的世界里，我们终于有机会发现这个错误，有机会更好地了解由我们一手打造的这个特定系统。通过这个bug，我们可以进一步地了解“技术实际上是如何运行的”，而不只是停留在“我们希望技术如何运行”的层面上。[17]

还有很多其他例子也表明了，虽然技术系统为意想不到的行为提供了便利，但同时也让我们有机会了解到技术实际上是怎样发挥作用的。例如，1982年初，温哥华证券交易所发布了自己的股票指数，其编制方法与标准普尔500指数（S&P 500），以及道琼斯工业平均指数类似。这个股票指数的初始点位定在1 000点，然而在接下来的两年内，点位却持续下跌。到了1983年底，点位已经跌至原来的一半左右。这样的下

跌是完全没有道理的，因为20世纪80年代早期的股票市场是一个大牛市。那么，这个指数为什么会持续下跌呢？有关方面展开了调查，最后发现这个指数的计算方法是错误的：它会直接把小数点3位之后的数字全部舍去。例如，如果计算结果为382.452 7，那么被存储下来的数值会是382.452，而不是四舍五入后的382.453。[18]

这个过程每天都要上演数千次，正确的数值无处可寻，这意味着损失了很大一部分价值。1983年11月，这个错误终于得到了纠正：前一周周五，指数收盘于500点左右，但是到周一开盘时，指数已超过1 000点，失去的点位重新回来了。这个计算过程中的深层问题，准确地说是算法中的错误，在这样一个异常现象中被注意到。事后看来，这起事件——当股票市场向好的时候，指数却在下跌，其实并不微妙。

我们还发现，企业和政府在利用复杂的机器学习算法，对海量数据进行挖掘时，也出现了不少类似的失败案例。在大多数情况下，这些机器学习系统对于公众，甚至对于专业人员来说，都是难以理解的“黑箱”。但是有时候，这些大型系统会表现出一些奇怪的或是令人担忧的“举动”，从而无意间为我们提供了探索系统内部世界的线索。

以微软的人工智能聊天机器人Tay为例。根据设计，这个机器人以19岁女子的身份与用户进行互动。微软将“她”放上了Twitter上，然而还不到一天的时间，“她”就已经“成长”为一名种族主义者。这是因为“她”所使用的算法接受了那些来自互联网的充满暴力的反馈信息。显然，Tay的偏执倾向并不是事先设定好的。通过这次失败，微软的设计师们更清楚地看到，程序是如何与原始的、未经过滤的互联网账号进行交互并结出恶果的，而这种恶果也是他们始料未及的。

显然，要想发现bug，“守株待兔”是远远不够的。许多技术开发人员都会主动搜索bug，并将它们集中放入数据库中，以系统化的方式加以解决。

更重要的是，在软件开发领域中，人们开始努力尝试打破这个怪圈。他们开始测试边界情况，搜集用户实际可能做出的各种怪事，而不再局限于预设固定的软件使用方式。例如，奈飞公司通过这种策略合理地开发出了“混沌猴”（Chaos Monkey）。[19]“混沌猴”的功能非常简单：它会出人意料地使奈飞系统服务中断。这里面的基本思想是：只有在观察到了庞大的奈飞系统如何应对各种故障之后，工程师们才能想出办法来维护系统的稳定，并以此抵御突如其来的各种意外情况。奈飞公司希望，在“混沌猴”完成了其使命后，工程师们所设计的系统运行方式，与

实际的系统运行方式能够完全匹配。

从失败中吸取教训，是理解任何复杂系统的重要机制。回顾一下科学史便不难发现，几个世纪以来，博物学家们一直在利用这种方法研究自然界中的复杂系统。

像生物学家一样思考

在我年纪尚轻时，对我而言，在所有词汇中，最令我感到困惑的一个词便是“杂项”（miscellaneous）。“杂项”这个词看起来是如此不可思议，我常常惊讶于它何以能够存在。这个词的有趣之处在于，它看上去像是用多个不同的语言单位拼凑而成的。我甚至不清楚它的准确发音是什么，但是它确实非常吸引人。

这个词的魔力不仅仅在于它的拼写和发音。看似混乱的拼写早已透露出它的含义：生活中，总有些地方是杂乱的。“杂项”意味着，即使是混乱无序，也可以构成一个类别或一种组织方式。“杂项”的存在本就是一种肯定：无论有多么不规则、多么杂乱，混乱的“秩序”是有可能被容忍和接受的。

接受“杂项”，即接受一种能容忍“杂乱”的精神，并不是每个人都能轻松做到的。在面对某个复杂情况时，包括我自己在内的许多人，本能反应是想要以某种方式去简化它，去除所有杂乱因素，并找出隐藏在背后的优雅结构。这种方法一旦成功，就会给我们带来极大的满足感。当我们找到导致失败的单一原因时，情况就是如此。但是，如果这种方法不起作用，我们就只能直面一片混沌，到那个时候，很多人都会感到不知所措。

与一般人不同，以研究世界“博物”为业的博物学家们，早就习惯了面对“杂项”。当他们发现正在观察的动植物的生活习性和行为习惯符合某种秩序时，他们会倍感欣慰；当然，即使不符合某种理论上的秩序，观察也会为他们带来其他收获。因为缺乏完整的理论框架，他们或许无法解释所见到的每一种生物，但是他们会记录下每一个细节，并尝试着去理解。有一个有趣的反例：某次，一位年轻人请物理学家恩利克·费米（Enrico Fermi）说出粒子物理学所研究的诸多粒子的名字，费米回答说：“年轻人！如果我能记住所有这些粒子的名字，我就是一名植物学家，而不是物理学家了。”[20]博物学家是什么样的？比如约翰·詹姆斯·奥杜邦（John James Audubon），他对美国境内的鸟类进行了分类和绘

图。与物理学家不同，博物学家认为应该了解每一个物种的细节，就算不知道所有物种之间的相互适应性，至少也要知道它们的名字，这是十分重要的。

在自然界中，也唯有通过研究生物进化中的错误和故障，譬如突变和疾病，我们才有可能了解生命系统的奥秘。例如基因复制过程中出现的错误，从染色体中的大型畸变，到脱氧核糖核酸（DNA）中不正确的代码复制，以及它们所导致的、可见的差异或缺陷，都是我们了解基因功能的突破口。研究果蝇的基因突变，有助于我们深入理解生命体如何从单细胞发育而来，以及基因蓝图如何培育出完整个体。具体来说，生物学家破解控制身体形态的关键基因序列的途径之一，就是观察一种可怕的触角足突变体，即一种在原本应该长触角的地方长出了腿的苍蝇。

对于技术系统，我们也需要采用同样的方法。“数字世界”是技术史学家乔治·戴森（George Dyson）(12)提出的一个术语，这个世界正在超越人类的控制范围。根据戴森收集的数据，在1953年3月，世界上只有53千字节的高速存储空间（RAM）。[21]时至今日，一台个人电脑所拥有的存储空间已是这个数字的10万倍有余。数字世界还在继续变得更丰富、更庞大、更互联化。它的发展速度是很多人都未曾想象到的。不仅如此，数字世界正变得越来越独立于人类。在全球范围内，信息的传播速度比我们能识别得更快，而且一直在以奇妙的、出人意料的方式相互缠绕着、作用着。

虽然无法理解所有技术系统之间的所有交互，但这并不妨碍我们成为“技术博物学家”。我们可以将系统及其各部分的多样性记录下来，并加以分类。即使无法完全理解整个系统，我们也可以通过考察异常情况和分析故障等方式，获得卓越的见识。

OVERCOMPLICATED

Technology at the limits of comprehension

复杂的技术系统更接近生物学系统，因此，用生物学思维思考复杂技术是个不错的选择。为了从整体上理解系统，我们也会忽略掉一些细节，这时，物理学思维才是首选。我们真正需要的是经过物理学思维锤炼的生物学思维。

05 为什么需要生物学思维

复杂的技术系统更接近生物学系统，因此，用生物学思维思考复杂技术是个不错的选择。为了从整体上理解系统，我们也会忽略掉一些细节，这时，物理学思维才是首选。我们真正需要的是经过物理学思维锤炼的生物学思维。

17世纪中叶，一位名叫纳撒内尔·费尔法克斯（Nathanael Fairfax）的英国医生在科学杂志《哲学学报》（*Philosophical Transactions*）上连续发表了几篇论文。[1]费尔法克斯观察到一些很有趣的现象，并决定通过这些论文把他的发现告诉同时代的科学家们。

其中一篇论文的标题是《人类自然特性潜移默化的实例：无论是在人身上还是在野蛮人身上》（*Divers Instances of Peculiarities of Nature, Both in Men and Brutes*）。在这篇论文中，费尔法克斯讲述了一个40岁左右、习惯喝热啤酒的男人的故事。有一天，这个男人在喝了一杯冰镇啤酒后病倒了，并在几天内死去。费尔法克斯就此推测：人的胃很可能只能适应一定范围内的温度。费尔法克斯还写了一个女人的故事：每次听到雷声，她都会觉得恶心。不过，费尔法克斯没有去推测为什么雷声会对这个女人产生这样的影响，只是指出："这位女士从孩提时代起就一直如此。"[2]

我们现在知道，尽管费尔法克斯没能实现他的目标——从记录下来

的这些观察结果和事实中总结出某种理论，但是他的观察本身就极具价值，至少这是“理解”的第一步。

在同一时期，年轻的牛顿正在思考物体如何移动，以及光线如何传播。牛顿于剑桥大学三一学院求学期间，一场瘟疫席卷了整个英国。为了预防疫情，剑桥大学临时闭校。于是，牛顿回到了家乡，在伍尔斯索普（Woolsthorpe）的农村生活了几年。[3]也就是在那个时期，他对微积分、光学和行星运动规律等方面的研究取得了根本性的进展。他研究了数学推理和演算，还做了一些实验，比如为了分析颜色的性质而在自己的眼窝上插入一枚长针，以及观察苹果如何从树上掉下来，等等。与费尔法克斯一样，牛顿对观察所得的材料进行了分类和编目；不同的是，牛顿还总结出了一套支配物理世界的定律，并用数学公式将它们描述了出来。

在某种意义上，我们可以说，费尔法克斯和牛顿在同一时期、同一个国度所进行的不同研究，代表了两种理解宇宙复杂性的方法，而这两种方法还存在着竞争关系。

牛顿试图将观察到的所有不同事物都统一起来，也就是通过一组优雅的解释来简化世界的多变性和多样性。通常，他所利用的只是几个公式或定律而已。关于这一点，我们可以在牛顿发现的万有引力定律中看得非常清楚。这个极简公式反映了从物体坠落、潮涨汐落到行星运行等诸多现象的普遍规律。而在今天，物理学家之所以会孜孜不倦地探寻能够一统天下的万用理论（Theory of Everything），也是出于与牛顿相同的愿景，希望能够发现可以作为人类已知的、宇宙各方面基础的秩序；并让宇宙的每个组成部分都各归其位，将它们放在适当的位置上。科学家托马斯·亨利·赫胥黎（Thomas Henry Huxley）有一句名言，“科学的巨大悲剧”[4]是“一个丑陋的事实往往会杀死一个美丽的假说”。他的意思是，优雅的理论是科学的目标，当某个事物与优雅的理论相悖，或令理论复杂化时，科学便会遭遇最大的悲剧。

费尔法克斯则放弃了对理论优雅性的追求，转而拥抱了多样性和复杂性。即使世界在一定程度上是混乱的，他也愿意接受，并为“又了解到新的细节”而欢欣鼓舞，哪怕这些细节很难立即被融入某个单一的理论框架中。有些人把他的方法戏称为“蝴蝶收藏家”式的方法，即收集多种多样的“蝴蝶”并加以描述。在这里，我们还能发现现代医生的影子，他们是费尔法克斯的智识后裔，为人体各个层级上的完美功能而啧啧惊叹，例如血液凝固过程中的复杂步骤、酶级联反应（enzyme cascades）的

复杂性质，等等；还有那些天文学家们，会为强大的太空望远镜所揭示的诸多星系类型而深深倾倒。

博物学家不可能赞同赫胥黎的抱怨，因为在他们看来，根本不存在所谓的“丑陋的事实”。所有的事实和知识都为我们提供了与这个奇妙世界有关的新信息，向我们展示着世界的复杂性和多样性。当事实不符合我们的心智模式时，完全不必为此而感到沮丧；相反，还应该为这种“意外”而感到由衷的高兴，然后去寻找能够解释这些“意外”的新方法。

物理学家弗里曼·戴森（Freeman Dyson）将牛顿的方法描述为古雅典时期的科学思维方式，并认为这种思维方式“强调思想和理论……试图找到可以将宇宙万物联系在一起的统一理论”。[5]至于关注多样性的方法，戴森则认为，可以将它描述为工业革命时代的科学思维方式，“强调事实和事物，试图探索和拓展人类对自然多样性的认知”。例如，起源于曼彻斯特的英国工业革命。

戴森还进一步指出：这两种方法还有另一个不同之处，即生物学是多样性理论者的“领地”，而物理学则是统一理论者的“主场”。我在这里把这两种方法分别称为生物学思维和物理学思维。在物理学中，人们通过统一和简化去观察各种现象的明显趋势，无论是在爱因斯坦、牛顿，还是麦克斯韦身上，都能看到这一点。众所周知，麦克斯韦给出了能解释电磁原理的公式。简化，甚至极简化，是物理学领域内广受尊崇的方法之一。

生物学家通常更愿意接受多样性，并倾向于陈列大量事实，而不在意这些事实是否能用某个统一理论来解释。[6]事实上，他们只需要有一个合适的小模型就可以了。当然，生物学理论也并非总是如此。例如达尔文的进化论，显然就是生物学中的一股统一力量；许多分子生物学家、以数理生物学为专业方向的应用数学家，以及许多研究其他领域的生物学家都倾心于此。

说到底，上述两种方法都是在探求具有普遍性的、有预测能力的理论。但是，这两种思维方式的推进方向是不同的，这种不同主要反映在它们对抽象化的相对容忍度上，而相对容忍度又取决于所研究系统的特性和复杂性。例如，利用数学公式抽象掉宏观层面上的细节，这种做法在物理学中几乎无处不在，但在生物学中却难觅踪影。

从下面这个古老的科学笑话中，我们可以清楚地看出这种区别。一位奶农为了提高产奶量雇用了两个顾问，一位是生物学家，另一位是物

理学家。生物学家在考察了一周后，提交了一份详细的长达300多页的报告，写明了每头奶牛的产奶量具体取决于什么，例如天气情况、奶牛的大小和品种等。而且这位生物学家还向奶农保证，只要严格按照建议执行，奶牛的平均产奶量可增加3%至5%。而物理学家只考察了3个小时就回来了，然后宣称自己已经找到了一个能够适用所有奶牛的高效解决方案，并且可以将产奶量提高50%以上。奶农问："那么，你说应该怎么做呢？""好吧，"那个物理学家回答道，"首先，假设你有一头身体为球形的奶牛……"[7]

抽象化方法当然是有用的，但我们不能做出存在"球形奶牛"这种假设。当你把生物学层面上的细节都抽象掉之后，你不仅会丢失大量信息，而且最终还会对某些重要组成部分，比如边界情况，感到束手无策。

生物学思维和物理学思维是解释世界的两种不同方法，适用于不同的系统，而且通常是互补的。

复杂的技术系统需要生物学思维的3个原因

那么，我们应该如何对待复杂的技术系统呢？它们是生物系统还是物理系统？理解技术系统应该采用哪种思维方式？现在我们就来探索生物系统和物理系统的特征，并与我们所了解的技术系统特征进行比较。

复杂的技术系统需要生物学思维有以下3个原因。

第一，生物系统通常比物理系统更复杂。

在物理世界中，系统的组成部分一般是相同的，也就是说，物理系统通常是这样的：由相同的气体分子组成封闭系统，或是由同一种原子组成单质，比如钻石。此外，在整个系统中，各部分之间的相互作用方式往往是统一的，例如卫星围绕行星运行的方式。

但是在生物学中，情况却不是这样的。在生物学中，系统的组成部分不仅类型繁多，而且涉及很多层级。例如，细胞中的蛋白质具有多样性；生物个体内部的组织器官也功能各异。要研究蓝鲸的交配行为，海洋生物学家就不得不考虑一切与之相关的事情，比如它们的脱氧核糖核酸，以及海水的温度等。生物系统中的每个组成部分不但各不相同，而且还很难从整体中被单独拆解出来。例如，要想观察一只变形虫的细胞核，并尝试分析它的特征和功能，就需要让细胞核留在生命体内，只有

这样才能了解细胞核如何适应变形虫的生活，如何提供涉及细胞诸多功能的核心遗传信息。现在，我们的技术系统正变得越来越复杂，显然，它们更像生物系统而不是物理系统。

第二，生物系统有别于物理系统的一个重要因素是，生物系统是有历史的。

生物会随着时间的推移而进化。尽管物理学所研究的对象也不是无中生有的，天体物理学家甚至还经常讨论恒星的进化进程，但是相比之下，生物系统更容易受到进化的影响。实际上，这正是生物系统的根本特性之一。生物系统之所以会呈现出如此复杂的结构，正是因为系统中存在极为复杂的历史路径，而且，在漫长的进化进程中，历史路径还会受到众多因素的影响。因为生命体的形式很复杂，所以任何微小的变化都有可能会带来意想不到的后果。随着时间的推移而发生的变化，一直在做着修补工作：以零敲碎打的方式修整着系统，使之适应新的环境。

举个例子，人类细胞中有许多至关重要的脱氧核糖核酸序列，譬如那些为遗传密码的翻译工作提供“指导”的序列，还有那些为使用能量提供“操作流程”的序列；而在其他生物体的细胞中，也存在同样功能的脱氧核糖核酸序列，尽管它们与人类的截然不同，甚至相隔千秋万代。这种生物学上的遗留代码有时可能是完全未经修改过的，但是通常来说，在进化进程中，生物系统也会被修补、被改变。

站在生命体的宏观角度上来看，这种情况意味着新功能通常是层层叠加在旧功能上的。这种分层有时会导致一些问题。例如，虽然人类现在已经可以双足直立行走了，但是人类的骨骼结构最初却是为了“四足运动”而存在的，也就是说，本来人类是用“用四条腿走路”的。后来，进化女神不断对我们祖先的身体修修补补，最后给我们提供了一个“足够好”的手段，也就是依靠脊柱实现双足直立行走。但是，这个解决方案远远谈不上完美，比如，有许多人都不得不忍受背痛之苦。

这种惊人的复杂性，以及对进化路径的依赖性，又会带来什么结果呢？在试图理解这些系统时，生物学家是不能依靠美学观念的。正如生物学家史蒂夫·本纳（Steven Benner）所说，进化而成的系统虽然能够有效地运行下去，但它看上去却不一定“很美”。[8]这样的系统可能很符合功利主义的需求，当然，有些人有时会把“功效”与“美”混为一谈，但是一般来说，它远远谈不上优雅或简单。本纳还进一步指出，真实的脱氧核糖核酸结构绝对算不上美，不过艺术家通过抽象手法呈现给公众的脱氧核糖核酸形象却很美。生物系统是进化而来的“足够好”的系统，而

不是设计出来的精致优雅的系统。换句话说，生物系统是拼凑而成的系统。

与技术系统一样，生物系统的进化也会留下遗留代码。[9]我曾在美国中西部地区居住过，当时我家后院种有许多皂荚树，树上密密麻麻地长着很多刺，那些刺又长又大，很是危险。[10]这些刺似乎没有任何功能，除了让我担心会被刺伤之外。皂荚树为什么会长刺？一个解释是，有了这些刺，这些树就可以“确保”叶子不会被猛犸象，或者其他北美巨型食草动物吃掉。当然，那些动物都已灭绝了。随着这些巨型动物的灭绝，对种子荚遗传基因进行编码的信息也势必会出现进化，当然，“长刺”这个信息到现在还没有消失。

虽然在这个问题上仍有很多争论，但许多科学家都认为，在生命体的基因组中存在着许多“多余物质”，这些物质也被称为“垃圾脱氧核糖核酸”片段。它们在进化进程中被积累起来并留存至今，却没有任何特定的生物学功能。许多复杂的技术系统，例如软件系统等，都包含着这类不会再被使用的功能，这些功能甚至可能已完全过时。与此类似，在许多生物系统中也存在不少这样的退化性状，它们原有的功能已不复存在。

对于生物系统和技术系统的相似之处，我们无法完美地一一对应。当然，生物系统处理遗留代码的方式与技术系统是有所不同的，[11]皂荚树最后可能会变得没有刺。如果这个性状确实是无用的，那么长刺对于皂荚树来说就是在浪费能量。因此，在未来的进化中，在与那些因为不长刺而更能适应环境的树种的竞争中，长刺的皂荚树会败下阵来，从而被淘汰。倘若真是如此，那么我的子孙后代就不用再承受被皂荚树刺伤的风险了，对此他们应该感恩。在大多数技术系统中，这种可能性都是不存在的。尽管遗留代码会让软件程序变得过时且效率低下，但它们却不会被自动清除。

第三，生物系统与技术系统的相似性还可以通过前文所提到过的高度最优化容限模型来进行分析。

某个看上去很强大的技术系统，可能会因为一些很小的干扰就出现灾难性的故障。在生物系统中也会发生同样的事情。例如，总的来说，人类具有很强的环境适应能力，但是人类基因组中的某个微小突变却会导致侏儒症，而且受到影响的两个基因副本是具有致死性的。无论是在生存尺度上，还是在构成材料上，人类以及其他生命体都与粒子加速器或计算机网络有所不同，但是所有这些系统的复杂性和脆弱性却有着深层意义上的相似之处。

总而言之，生物系统与技术系统确实存在着深远的“亲缘”关系，这意味着我们可以从生物学思维中学到很多东西。

技术领域的“生物学家”

技术系统变得越来越复杂了，我们对它们的理解也走向了两个极端：要么，只能得到关于系统运行的一般性概念，但对其内部细节印象模糊，甚至一无所知；要么，对系统的若干组成部分有零碎的了解，但并不知道这些组成部分是如何融合在一起的，也不知道该对系统行为作何预期。前者趋于物理学思维，而后者则趋于生物学思维。

面对越来越强的复杂性，许多人选择通过物理学方法，抽象掉细节来获得对系统的一般性认知。例如，在考察一个复杂的社会系统，比如一家大公司或一个城市时，如果用物理学思维来解释，那么我们所采取的方法可能是：将它的其中一个属性绘制成图表，然后看图表能否很好地拟合为一条特定的数学曲线。这种方法能让我们清楚地看到系统内部所发生的事情，至少可以提供相关的线索。但是，这个系统及其行为之所以能很好地与某条曲线相拟合，原因可能有很多。在这种情况下，你得到的只会是更多的问题，而不是答案。理解复杂系统通常不适合采用直接的、大规模的抽象化方法，因为这些系统实在太混乱、太庞杂了。

因此，注重细节、强调多样性的生物学思维，为理解杂乱的进化系统提供了一个至关重要的视角。只有通过大量的刺激和检验，复杂的进化系统才有可能被完全理解。因此，生物学家，特别是野外生物学家，在研究生命体的复杂性和多样性时，都必定会考虑它们的进化轨迹。这种方法特别适合用来理解技术系统。野外生物学家经常要扮演博物学家的角色，非常注意收集和记录在各地发现的各种事物，并加以分类和编目。除此之外，在面对一个极其复杂的生态系统时，野外生物学家也不会即刻就想要彻底理解它。他们很清楚，一次只能研究这个系统的一小部分，而且即使是针对某一部分的研究，也未必会完美无缺。例如，他们只研究少数物种之间的相互作用，鲜少会去考察某个地区的完整的物种网络。野外生物学家对自己的判断非常坚定，而且明白，在任何时刻都只能观察周遭复杂情况的一个片段。

与此类似，在面对一团乱麻般的技术系统，譬如一个软件、某个国家的法律体系，或是整个互联网时，如果我们硬要将物理学思维中的优雅和简洁附加在其整体性之上，那么我们就不可能走太远。要想从真正

意义上理解技术系统，并对其行为做出有效的预测，我们就需要成为技术领域的野外生物学家[12]。

这对我们意味着什么？在思考一个系统的不同交互层级时，我们要记住，那些看上去毫不起眼的底层细节有可能会升至顶层，从而变得对整个系统至关重要。我们需要技术领域的野外生物学家，需要他们来对复杂系统的细节和各个组成部分及其失败和错误进行编目和研究。这种生物学思维不仅可以带来新的思想，而且完全有可能成为我们探索互联性越来越强，而可理解性越来越弱的技术世界的主要路径。

如前所述，我们可以在技术系统的错误中学习，就像生物学家在遗传错误中进步一样。但是，生物学家所做的并不仅限于此。为了更好地向生物学家靠拢，我们必须更加深入地了解生物学的一般研究方法。

近年来，遗传学领域的一项重大突破是核糖核酸干扰技术（RNA interference）。核糖核酸是脱氧核糖核酸的“表亲”，在细胞中生成，并服务于多种目的。这项技术的核心是，利用核糖核酸的小片段终止某些蛋白质的合成。这就是说，在构建出正确的核糖核酸文本之后，你就可以有效地关闭某些基因了。

那么，这个机制是怎么被发现的呢？最初，一家初创的生物技术公司的遗传学家们，试图用基因工程方法培育颜色更深的紫色矮牵牛花。[13]他们在识别出了矮牵牛花中负责显示紫色的基因后，便设想再加入另一个基因拷贝，以使花朵颜色更浓艳。然而，当他们真的这样做了之后，得到的却是一朵白花，这与他们的设想背道而驰。新加入的遗传信息不但未能让紫色矮牵牛花变得更浓艳，反而使它完全失去了紫色。不过，研究者们既没有因这个意想不到的结果而感到气馁，也没有忽视这个结果，他们敏感地意识到，研究需要继续。最终，核糖核酸干扰技术诞生了。

事实上，这种事情不仅常见于生物领域，在其他科学领域也经常发生。美国著名作家、生化学家艾萨克·阿西莫夫（Isaac Asimov）曾指出：“在科学研究中最激动人心的短语，也是预示着新发现的短语，并不是‘Eureka！’（有了！找到了！），而是‘That's funny…’（有趣的是……）。”[14]生物化学家亚历山大·弗莱明（Alexander Fleming）在培养皿中看到了一些预料之外的奇怪东西，他没有放过它们，于是发现了青霉素。原子核也是科学家们在检验一个实验的意外结果时发现的，他们原本是想用一片薄薄的金箔映射放射性粒子。科学家们没有对意外结果视而不见，反而通过仔细观察，得出了对原子结构的全新理解。如果

借鉴博物学思维，对那些看似不合理的少数特例进行收集和编目，并养成习惯，我们就能对正在研究的事物拥有全新的认识。

当然，仅仅被动等待意外发生，然后加以观察，是远远不够的。生物学家还会积极地采取行动，主动将意外事件注入系统，然后观察系统的反应。在试图培养某种特定类型的细菌，例如某个可能会生成特定化学物质的细菌变体时，他们会采用一种通常被称为“诱变”（mutagenesis）的方法。顾名思义，诱变就是通过积极的尝试来诱导产生突变。例如，对生命体进行辐射，或让它们接触有毒的化学物质等。

虽然这种做法听起来会令人稍感不适，但它绝不是无的放矢。当我们预测不出某个复杂系统可能会产生何种反应时，或者无法确定基因组中的哪些变化可能会产生预期的效果时，我们通常需要利用一定的随机性来找出系统的行为倾向。从根本上说，这些系统都是非常复杂的高度非线性系统，所以我们不得不“借用”自然进化进程中的修补措施，来探索它们的运行方式。

在制药公司研发新药的过程中，也可以总结出类似的工作原理。在研制新药时，有一种方法是：对一种化学物质的多种变体进行筛选和试验，以找到符合疗效的化合物。[15]虽然还无法完全理解化学机制的“真相”，但这样的试验确实帮助我们找到了一些不可多得又行之有效的药物，并让我们对人体有了更深刻的理解，尽管这种理解有时候是间接的、模糊的。为了能更多地了解并筛选出可以合成有效药物的分子结构，我们必须不停地“拨弄”系统。制药业的经验已经证明，这种方法不仅有效，而且是科学发现所需要的修补方法之一。

因此，要将生物学思维应用到对技术系统的研究中，我们就必须认识到，“修修补补”是构建系统和理解系统的一种重要方法。正如斯图尔特·布兰德在讨论遗留系统时所指出的那样：“要从遗留系统中汲取新功能，不是简单地下一个命令就能实现的，而要通过一系列谨慎的实验，在‘运气’的帮助下，才有可能得到我们想要的结果。”[16]这就是技术领域的生物学方法。上一章中提到过奈飞公司的“混沌猴”，本质上即是一种诱变软件，旨在通过引入故障来更深刻地理解系统，并提高其稳定性。对于所有故障和问题，包括人为引入的，我们都必须记录下来，并进行仔细分析，以更好地理解大型系统的运行情况。

这种生物学思维还可以帮助我们更好地理解天灾人祸。比如说，我们需要借鉴生物学家对癌症的观点：在某些细胞长成肿瘤后，不能简单

地说，这是某件事情出了问题。癌症其实是诸多因素和多种生物反应累积所致，而且在这些因素和反应之间还存在复杂的相互作用。癌症是身体出现的大规模故障，这种故障足以致命。与此类似，对于技术系统，我们也要认识到，问题以及系统对问题所做出的各种反应，在积累到一定程度时，可能会导致同样的级联式失败。这就意味着，我们必须像生物学家那样思考问题。在一座核电站中，不断累积的细小问题最终会引发严重问题。例如，在三哩岛核电站事故中，其实有多个独立因素都与该核电站的部分熔毁有关联。[17]同样，我们可以用这种方法将导致金融市场崩溃的各种因素之间的相互作用识别出来。

令人高兴的是，我们为理解技术系统而做出的努力，已经得到技术系统自身的帮助了。在今天，有很多计算工具都可以帮助我们发现系统中的意外结果。这些计算工具皆来自“新颖性检测”领域。也就是说，现在的机器已经可以与技术领域的野外生物学家和博物学家合作了，它们可以帮助我们更好地理解，至少是部分理解技术系统。

当物理学遇见生物学

如前所述，技术具有诸多“生物学性质”。它不仅“笨拙”，而且会在进化进程中因为修补而生长变化，同时还拥有许多繁杂的细节。那么，这是不是意味着，我们应该放弃探寻复杂性背后的规律呢？绝对不是！在理解技术系统时，物理学思维也能发挥重要作用。

我们在试图理解一个复杂系统时，必须先确定以何种分辨率来考察它，或者说，必须先确定所考察的细节涉及哪个层级。在我们所关注的层级上，细节的精准度如何？应该关注生命体细胞中的单个酶分子，还是身体内部的器官和血管？应该关注通过电路传播的二进制信号，还是着重测试计算机程序的整体结构和功能？在更高的维度上，我们是否只需要关注计算机网络的一般属性，而不用关注组成这个结构的单个机器及其决策？

这些问题并不总是很容易回答。有时，我们必须向物理学靠拢，通过抽象掉部分细节来理解整个系统；有时，细节又特别重要，譬如我们在前文中讨论过的罕见词和边界情况，在这种情况下，我们必须更多地依赖生物学思维。

然而，在很多情况下，不同水平的分辨率是会产生冲突的。镰状细胞性贫血是一种相当严重的全身性疾病，但它的致病因素仅仅只是脱氧

核糖核酸中单个碱基对的一个微小变化。发生于2003年夏天的俄亥俄州电网崩溃事件表明，即使只是一棵树碰到了电线，也能引发级联效应，导致全美电网大面积瘫痪。不同系统之间的联系变得越来越紧密了，在这种情况下，不但不同水平的分辨率会相互交织，而且原本相对独立的各个领域也会越来越多地产生交集。这就意味着，我们必须越来越多地把物理学思维和生物学思维结合起来，在探寻秩序的同时，也不能忽略粗糙的边缘。在将生物学思维与物理学思维结合到一起后，我们便能更加了解周围各种拼凑起来的系统了。尼尔·斯蒂芬森（Neal Stephenson）在其小说《编码宝典》（*Cryptonomicon*）中借人物之口，详细地描述了希腊神殿的结构。书中人物所说的话与我们在此阐述的思想不谋而合：

> 然而，正是这种混乱、这种不对称性，才使万神殿更加可信。就像元素周期表和基本粒子系谱一样，也像你通过尸体解剖看到的人体结构一样，它除了呈现出足够可靠的、我们的大脑能够分析的模式之外，还潜藏着一种不规则性。这种不规则性表明，“有机来源”是存在的。例如，这里的太阳神和月亮女神是明确对称的；这里还有赫拉（Hera），她并不在这种对称性中扮演角色，单纯只是一个女神；然后还有狄俄尼索斯（Dionysus），他甚至不完全是神，而是半人半神，无论如何，他也进入了万神殿，与其他众神一起坐在了奥林匹斯山上。这就好比，你走进最高法院后却发现一个马戏团小丑坐在了一群大法官中间。[18]

我们对身边的各种系统审视得越是仔细，就越能看清生物学与物理学之间的平衡关系。在生态系统中，在我们每天依赖的混沌技术中，皆是如此；在希腊神话中，在讲给自己听的故事中，亦复如是。

事实上，“讲故事”是一个非常好的方法，能够让我们融汇生物学思维和物理学思维。有些故事就像精心设计制造的机器，毫无赘述，所有情节都非常合理。这方面的例子最经典的莫过于“契诃夫之枪”原则，这个原则源于伟大的剧作家安东·契诃夫（Anton Chekhov）的创作理念：故事中引入的任何元素都必须与剧情有关。例如，在第一场中出现过的一支上了膛的枪，在第三场中必须开一次火。

当然，也有这样一类故事：感情丰富，铺陈繁多，所出现的元素不一定会推动剧情发展。例如，荷马在《伊利亚特》中叙述的入侵伊利亚特的希腊战船；克莱默（Kramer）在《宋飞正传》（*Seinfeld*）中提到

的从未露过面的宋飞的朋友鲍勃·萨卡马诺（Bob　Sacamano）、洛米兹（Lomez）、科尔基·拉米雷斯（Corky　Ramirez）等。虽然不是推动故事情节发展的关键元素，但它们同样很重要。[19]这就是与物理学肩并肩的生物学，当我们讲故事时，在增添故事丰富性方面，两者皆不可或缺。

特效领域有一个耐人寻味的术语："小东西"（Greeblies）[20]。每当听到这个术语，我就会想起小精灵，或者伶牙俐齿的小丑。"小东西"指的是被添加到某个场景中或某个对象上的小物件，目的是让场景或对象看起来更可信。这就好比，你制造了一艘未来主义风格的飞船，那么总不可能让它完全由一些转角和若干光滑的面板组成吧；它还需要加上出入口、通风口、各种装饰物、管道、隆起物，以及凹凸槽等。《银河战星》（*Battlestar Galactica*）或《星球大战》等影片中的宇宙飞船，之所以拥有强烈的吸引力和视觉冲击力，就是因为它们都拥有这种"莫名其妙"的复杂性。

对于为什么要不断加入小物件，著名数学家、"分形"的创造者本华·曼德博（Benoit　Mandelbrot）的一段名言给了我们很好的解释："为什么有人会说那些基本的几何形状是'冷冰冰的'和'干巴巴的'？其中一个原因就在于，它们无法被用来描述云、山、海岸线或树的形状。云不是球形；山不是圆锥体；海岸线不是圆形；树皮不是光滑的平面；雷电也不是按直线行进的。"[21]

技术系统也是如此。一旦嵌入了真实世界，技术系统就失去了如直线、三角形般纯净的逻辑结构；相反，随着时间流逝，它将因吸积而充满各种杂项，就像生物系统中因积累而出现的"进化大杂烩"一样。故事要以细节和复杂性为基础，系统也一样。事实上，我们之所以会认为某些东西更加"现实"，正是因为它们是复杂的，充满了细微的"锯齿"和各种细节，尽管这些"锯齿"和细节常常是我们一时无法理解的。说到底，我们需要这种混乱，需要这些小物件，即使它们在分辨率水平较低的解决方案中会被全部抽象掉。

生物学思维必须与物理学思维"和谐共处"。不妨回想一下博尔赫斯的短篇小说《博闻强记的富内斯》，主人公富内斯背负着完美记忆的重担。[22]虽然许多人可能会像富内斯一样，将这种能力视为一份天赐的礼物，但事实并非如此。我们可以从小说中得知，对富内斯来说，几乎每一个细节、每一种观点都会形成新的记忆，继而形成新的记忆类别或类型。当他凝视一只狗的时候，他无法像普通人那般看待狗；每时每刻，

他都是在用某个特定视角看待一只特定的狗，只要狗，或他自己动了一下，那么他的大脑就会生出一个关于狗的全新记忆。富内斯没有办法将所有细节统一到一起，并形成某个一般性概念，因为他的记忆过于细节化。换句话说，他无法形成任何抽象的概念，因为他对庞杂现象的容纳度实在太高了。

事实上，生物学家的最终目标同样是创建模型，并识别出规律，只不过这个目标所涉及的范围可能较小。在面对一项复杂技术时，我们需要从野外生物学家的角度出发，围绕其边缘进行各种各样的实验，看看它会做出怎样的应对。当然，在这样做的时候，我们的最终目标仍然是求得某种程度上的一般化。这种方法并不鲜见。很多人在玩诸如《我的世界》（*Minecraft*）这类开放式游戏时，所采用的便是这种方法。在这类游戏中，你先要收集大量与虚拟世界有关的信息，例如，可以做什么，不能做什么，什么东西会“杀死”你，怎样才能生存下去，等等。然后，你需要构建一个比较小的心理模型，接着在较大范围内求得一般化，当然，相对于整个世界，那仍然只是小范围的一般化。

又例如，你正在使用一个很高级的计算机软件，比如说一个庞大而繁杂的文字编辑软件，然后你发现文档尾注的编号变得凌乱了，一般来说，你并不会因此而感到惊慌失措。你会先看看到底出了什么问题？有几个尾注的编号重复了？它们还能关联到文档中的正确位置吗？通过这种修修补补的生物学技巧，你会对这个复杂系统的细节了然于胸。只要真正地像技术领域的生物学家那样行事，我们就能做到，在对系统各个部分进行仔细研究的同时，还能清醒地认识到它们只是一个更大的、高度互联的整体的组成部分。

接下来，我们的讨论将转向另一个领域，其间的知识可以帮助我们在把握技术系统的同时，实现微妙的平衡。

复杂性科学的视角

复杂性科学，是管理和理解复杂系统的一条自然路径。它能够定量地研究庞大且复杂的互联系统，无论是生物个体，还是生态系统；大到万维网，小到电影演员的合作网络。提到后者，我们自然会想到凯文·贝肯定律，也就是六度分隔理论。复杂性科学是一个发展迅猛的、令人振奋的研究领域，我自己也涉足其中。它利用各种强大的思想和数学框架，探寻着复杂系统的模式和含义。它所采用的方法极其广泛，从理解

网络结构，到构建拥有大量交互主体的计算机模型，也就是基于主体的模型（agent-based model）。

不过，这种理解系统的方法需要先抽象掉一定数量的混乱因素，并找出规律，以进行清晰的数学推理和分析。如果你正在观察一个巨大的交互网络，比如人们在Twitter上的互动，那么你可以先忽略交互的内容和其他细节，只关注个体之间的关联数量，你会发现，关联数量可以被拟合成一种特定类别的概率曲线，即重尾分布（heavy-tailed distribution）[13]。这个曲线不会告诉我们谁与谁有关联，谁的关联数量很多，谁的关联数量很少，但它足以表明，交互网系统中存在基本的统计规律。

在复杂性科学中，还有一些模型探索的是所谓的渗流（percolation）：流体物质如何基于某种多孔结构进行扩散，例如，石油如何在岩石中渗透。这些模型能够有效地说明，密度的微小变化会导致流体物质在系统中的扩散能力发生巨变，尽管这些模型无法向我们提供诸如“孔隙的确切位置”之类的信息。

毫不夸张地说，复杂性科学中的每个模型都为理解系统提供了独特的视角。事实上，很多时候我们都是在借助简单的数学模型，统一着世间的诸多事物。例如扩散限制凝聚模型（diffusion-limited aggregation model），这个模型表面看来不过是一个抽象的玩物，但是实际上，它可以被视为多种现实现象的统一模型。

扩散限制凝聚模型简称为DLA模型，其原理是：先在巨型网格的中心设定一个点，然后让其他点在网格上随机移动；在某个点击中初始点之后，那两个点就形成了一个聚合体；接下来，让剩余的点继续移动，直到它们中的某个点又击中聚合体。因为加入聚合体的点越来越多，所以聚合体缓慢地“长大”了。

聚合体最终会变成一个形状怪异的东西，不能被称为“团”（blob），也不能被称为“畸形的圈”，也不是普通的结晶体，而是一种更加“有机”的东西，如同生物中的徒长（spindly growth）[14]。事实上，它的“模样”取决于模型的类型，有时候会像珊瑚，有时候会像闪电，甚至会像城市。由此可见，一个简单的、无足轻重的计算机模型，可以生成非常复杂的图形图像。

这个模型，以及其他一些类似的模型，将一个“悖论”带到了我们面前：我们可以通过非常简单的模型来复制各种复杂系统，当然，具体细

节是不一样的。这些小型“玩具模型”具有某种“波特金复杂性”（Potemkin complexity）。“波特金复杂性”是指，某些系统看上去很复杂，但本质上却很简单。例如，由它们复制而成的“大城市”与真正的大城市有所不同；真正的大城市，当前的状态包含着大量以往的决策。由它们复制而成的“珊瑚”也与真正的珊瑚不同，后者的所有细节都会融入它的生长过程。

尽管如此，在很多时候，有这些模型就已足够。对于某些基于特定目的的观察，例如，在试图理解一个实体组织的组织形式时，我们所关注的就是总体形态。对复杂系统的抽象化和简化，不仅可以帮助我们理解系统行为的总体情况，还可以告诉我们系统的反馈机制、互联性，以及对初始条件的深刻依赖性，甚至可以告诉我们系统会对变化做何反应，并在此基础上帮助我们设置边界。利用复杂性科学的相关知识对复杂系统进行高效思考，是当下每个人都应该掌握的技能。不管是专家还是普通人，只有具备了这样的技能，才能在这个日益复杂的世界中好好地生存下去。因此，教会人们如何把握复杂系统的运行机制，成了一件迫在眉睫的重要任务。

但是，复杂性科学的一般化模型也是具有局限性的。在试图了解不同系统的特殊情况时，简单模型就不够用了。在这一点上，菲利普·鲍尔说得非常到位：“河道网络的互联模式与视网膜神经的互联模式，既有相同的地方，又有不同的地方。将它们统称为分形模式是不够准确的，它们各自的分形维数也是不够精准的。要想充分地理解河道网络，就必须考虑物质的沉积和输送、气象条件的变化、河道底层基岩地质的具体性质和差异等一系列情况；而这样的思考过程与神经细胞网络毫无关系。”[23]

不过，令人欣慰的是，复杂性科学所提供的工具不仅可以用于创建经过简化的抽象模型，还可以为我们探究系统细节提供窗口。也就是说，它可以积极地、严格地筛选出系统的复杂性。举例来说，一些研究者利用软件工程和复杂性科学的某些方法，分析出了《美国法典》这个法律系统的诸多性质。[24]这项研究的分析结果体现出了这个法律系统的很多总体特征：其不同组成部分的复杂程度不同；在总体上存在着一定程度的庞杂性，等等。同时，他们还筛选出了特别有意思或特别错综复杂的一部分法律，比如拜占庭式的、有隐秘含义的法律。在涉及银行义务的那部分条文中，突然跳出了详细说明“公司的权力和义务”的一节，而且内容特别复杂，至少从条件的数量上来说是如此，它的存在就像是

软件系统中的“if-then”语句。与此类似，在关于税收的部分，“适格的养老金、利润分配和股票红利计划”一节也特别复杂。这些研究结果显然有助于我们集中关注，并“放大”分析法律体系中最复杂的部分，甚至有可能会起到促进简化的作用。

与野外生物学方法一样，某些复杂性科学方法可能更适合用来了解复杂系统中某个子系统的行为，而不是整个系统的行为。我们还可以利用复杂性科学的方法来找出异常值，也就是复杂系统中不符合一般规则的部分。在所有这些方法的共同作用下，我们便可以确定系统中哪些部分更值得深入研究，从而更好地了解它们是如何相互作用的。我们必须协调好抽象化的复杂性科学方法与其他方法之间的关系，后者往往可以帮助我们找到“无法平滑嵌入模型”的特定部分和细节。

在这里，我们必须再一次面对物理学思维与生物学思维之间的紧张关系。一方面，我们希望技术世界拥有简单和优雅，希望那些混沌不明的东西消失；但是另一方面，我们必须承认，有些事物既是庞杂的，又是微妙难解的，还会随着时间的推移发展和进化，这便是心智成熟的标志。在成长为独立的个体之后，我们会认识到人际关系的复杂性，认识到人际交往的很多精致微妙之处。与此类似，当我们所处的社会逐渐成熟起来之后，人们必定会认识到社会结构中固有的复杂性和不规则性。复杂性科学，尽管绝非“万妙灵丹”，但确实可以帮助我们实现这种平衡：既能突出需要特别关注的细节，又能为知识范畴和关注程度设定恰当的边界。例如，复杂性科学可以告诉我们，系统为什么很容易就会变得不稳定：如何变得不稳定，以及我们必须特别关注哪些地方。如果一个简单模型已经证明了，某个大型技术网络中的一个微小变化很可能会导致网络彻底崩溃，那么我们绝不能对此保持“无知的幸福感”。

知易行难。想要在物理学思维与生物学思维之间找到平衡点，并非易事。这是一个艰难的探索过程。在漫长的思想史上，人类一直沿着这个方向不懈地探索着。

思维方式的进化

事实上，许多先贤都曾试图理解周遭世界。在历史上，爱奥尼亚（Ionia）是最早出现这种严谨思想活动的地区。爱奥尼亚是一批古代城邦国家的总称，这些国家位于爱琴海之畔，也就是现在土耳其的西海岸，古希腊的第一批哲学家就出现在那里。在苏格拉底时代到来之前，

他们已开始潜心思考宇宙的奥妙。这些古希腊哲学家的思想是前苏格兰学派思想的一部分，并且拥有一致的特点：压倒性的简化和一般化。[25]

爱奥尼亚诗人兼哲学家色诺芬尼（Xenophanes）说过两件事，或者说是一件事，这取决于你的理解：世间万物，皆源于“土”和“水”；归根到底源于“土”。无论如何，色诺芬尼是在尝试着用一个或两个基本原则来解释世界的复杂性。

与此同时，另一位爱奥尼亚哲学家泰勒斯（Thales）则认为，万物皆由“水”组成；还有一位哲学家认为，一切皆来源于“火”。这样看来，“地球超人”（Captain Planet）(15)应该会喜爱这些爱奥尼亚哲学家。

米利都的阿那克西曼德（Anaximander of Miletus）是又一位爱奥尼亚哲学家，在他看来，一切事物都源于单一的、无限的第一原则。阿那克西曼德还提出了其他许多颇为精确的理论，例如，动物来自“被太阳蒸发的水分”。[26]他还假设，地球正被一个相当于自身体积28倍大的圆圈包围着，并由此推导出了不少天文学理论。

不过，总的来说，这些哲学家都倾向于：只需少数几种基本材料和少数几个基本原则就可以解释整个世界。或者说，他们几乎都坚持了古希腊人关于kosmos（宇宙）的基本观念。与现代英语中的cosmos（宇宙）有所不同，希腊人不仅用这个词来指代宇宙及万物所依赖的整体系统，还用这个词来形容宇宙的优雅和美丽，以及终极秩序。也就是说，大自然的优雅和保持秩序的物理定律，其实都已隐含在这个表达自然的词汇之中了。除了“宇宙”这个术语之外，这些哲学家还常常使用另一个术语：“本原”（arche），其含义与作为宇宙基础的指导规则，也就是宇宙法则很是接近。很显然，这是一种渴望统一世界秩序的愿景。[27]

离开古希腊，来到现代社会的黎明时代，我们会发现思想家们的观点发生了一些转变，他们更加关注奇怪的、无法解释的事物和现象。我们看到了“杂项”。根据菲利普·鲍尔的研究，古代和中世纪的学者对异常现象和古怪行为的兴趣普遍不高，他们感兴趣的是确认已知的东西。[28]但是，对于那些致力于建立“珍奇博物馆”，站在现代社会入口，处于风口浪尖上的开创者们来说，那些异常现象和古怪行为才是这个世界上最令人激动的事物。科学史学家洛琳·达斯顿（Lorraine Daston）认为，在科学发展的早期阶段，也就是在费尔法克斯时代到来之前，科学家们都醉心于“奇怪的事实”。[29]

关于这一点，可以来看看达斯顿所引述的，于17世纪发表的一系列科学论文的标题：《怪物头上的可见物》《背上有洞的野猪，当它被猎人追赶时，洞就会冒泡》《一位潜水员的讲述：恐怖的雷电奇闻》……从这些标题中，我们不但可以发现不少怪异的信息，也不难判断出当时的科学家喜欢收集奇异事物。他们渴望记录和解释自然世界中那些最意外和最古怪之处。正如达斯顿所指出的那样："第一批科学事实是很难处理的，并不是因为它们很稳定，我们无法对它们视而不见，而是因为它们太怪异，我们无法将它们归于某个理论。"[30]

很难想象，现代科学作为收集新奇事物的手段之一，却没有办法解释那些新奇的东西，但这就是当时的科学，存在于生活的方方面面。当今时代，我们不再把科学视为收集古怪事物的一种活动了，更不会认为科学就是对这些事物的解释。然而在历史上，科学的确曾是这样的活动，而且直到今天，在很多方面仍有这样的"遗风"。当我们尝试认真面对环境的畸变和异常现象时，我们通常能够学到更多的东西。例如，物理学实验中的一个异常现象最终揭开了原子核结构的面纱；核糖核酸干扰技术的诞生也可以归因于一次意外。

这是一种融合了物理学思维的生物学思维。我们必须把这种思维方式引入到技术领域中。对于这项工作，从一定程度上来说，人人都可以参与其中。此外，我们还需要技术领域的野外生物学家，需要他们在复杂的技术大厦中挖掘"意外之物"，并与难以理解的事物"做游戏"。那么，这些人会是谁？在当下这个专业化时代，我们需要一些通才，他们有能力识别并牢牢抓住系统的各种细节，以及未被理解的边界情况和关键节点。

我们需要通才

2009年，一个科学家小组利用在线科学论文网站，对用户的数亿次互动进行了分析，希望能辨析出读者的"点击流"，也就是读者从一篇论文出发，去往下一篇论文的路径。[31]研究数据最终揭示了人们从一个主题领域移动到另一个主题领域的行为模式。科学家甚至还生成了这些主题领域的互联"地图"，图像非常漂亮。在这个"地图"中，护理学位于心理学和教育学附近；有机化学是物理化学和分析化学之间的纽带；无论是经济学还是社会学，都和法学有关联；音乐则显得有点"特立独行"。

当然，这样的互联"地图"有些过于简单了。实际上，音乐也会涉及

物理学和心理学的概念，而经济学则会大量地从数学中汲取营养。不过，无论如何，分析“点击流”确实是探索思想互联性的有效途径之一。为了理解我们自己所构建的、越来越复杂的系统，专业化分工仍在继续深化，但是互联“地图”提醒我们，没有哪个领域可以孤立存在，每一个领域都是这个广泛互联的系统的一部分。

因为各种系统会以多种不同的方式相互关联，所以人类越来越需要将某个知识领域与另一个知识领域联系起来。例如，在编写游戏《危险边缘》（*Jeopardy!*）的计算机程序时，你需要运用从语言学到计算机硬件领域的相关知识。在这种情况下，仅靠专业化是行不通的，我们还需要依赖相当广泛的知识。但是，如前所述，我们的知识体量很快就会触及上限。是的，人类大脑的生理特性决定了人们无法掌握所有知识。

为了克服这种困难，我们需要培养通才。在这里，通才是指既能看到“土地的地形地貌”，即拥有抽象的物理学思维；又能在尚未理解整个系统的情况下便懂得欣赏系统细节，也就是拥有复杂的生物学思维的人。毫无疑问，在构建复杂系统时，通才是最适合成为博物学家和野外生物学家的人，而且必不可少。他们可以从一个片段跳到另一个片段，检查出那些没有意义的部分，并在巨大的技术系统中挖掘出正在发生事件的线索。

或许有人会问：既然知识量增长得如此迅猛，那么当今世界是否还有通才存在？

当今世界，通才当然是存在的，只不过想要成长起来却非常艰难。要培养出能够很好地扮演上述角色的通才，第一步是要先培养出一大批“T型人才”。这个术语最早出现在计算机教育领域，[32]而后泛指在某个领域内既拥有高深的专业知识，也就是拥有T型中的“主干”那一竖；又拥有广泛知识，也就是拥有T型中的那一横的人。

那么，这种类型的人应该是什么样的呢？数据科学家正是T型人才的典范。数据科学家的职责是利用计算机科学和统计工具挖掘大型数据集中的隐藏含义。这项工作与具体学科没什么关系。为了很好地完成工作，数据科学家必须了解多个领域的专业知识。我们在应用数学家身上也看到了类似的特征，他们要跨越学科界限，用定量工具找出各学科的共同之处。这便是通才的模样。

通才还可能出现在咨询和图书编辑等行业中，你甚至可以在风险投资界看到他们的身影，在那里有很多人不但拥有多领域的丰富知识，而

且还懂得如何有效利用这些专业知识。如果有人想同时在外层空间、3D印刷、农业技术和科学探索工具等领域有所斩获，那么他必须得拥有通才的潜质。

通才既拥有专业知识，又乐于探索诸多不同的领域。但是，要想培养出T型人才并不是一件简单的事。他们先要将专业化与普遍性结合起来，才能着手去解决那些日益复杂的问题。

在当今社会中，专业化人才深受就业市场的追捧，这让培养T型人才这件事难上加难。[33]在我看来，在未来相当长的一段时期内，我们都很难在学术界找到这类才人。正如商学教授大卫·蒂斯（David Teece）所指出的那样，只有在文艺复兴时期，成为“文艺复兴人”才那么容易。但是，从培养“女童子军”这件事中，我们看到了一丝曙光。[34]

女童子军组织曾颁发过一种非常有意思的荣誉徽章：“业余家”徽章（Dabbler badge）。此举是为了对那些想在每一件事情上都做好的女童子军们给予认可与鼓励。这种做法是值得推广的。如果你不想只学制陶或摄影，而想多学一些有用的东西，并且你样样都学得不错，那么“业余家”徽章即是为你量身定做。我的母亲曾经告诉我，她当年在参加女童子军后，获得的唯一一个徽章就是“业余家”徽章。

或许，现在是时候推广“业余家”徽章了。对那些致力于了解各领域思想和技术的人，我们应该就其能力和贡献给予认可和鼓励，无论是在其小学阶段，还是在其职业生涯的最后阶段。他们可以帮助行业专家们将一个领域的专业知识“翻译”到另一个领域中去，以此化解“隔行如隔山”的困境，抵御“不可理解性的洪流”。这些“业余家”可以为巨型系统的构建工作带来生物学思维：找出系统中可能会发生的故障，对这些系统做出深入的理解，并对系统中那些奇异和惊人的组成部分进行分类和编目。

事实上，最令人费解之处往往最能让通才感到如鱼得水。在那些令人费解的地方，系统是如此复杂、如此紧密地相互联系着，以致我们只能尽力将“杂项”记录下来。这就意味着，通才的培养不仅在于学习已知知识，更在于探索未知领域，寻找前人所不及的新方法。而这又意味着，通才不仅要掌握严谨的人文科学思维，还要拥有数学思维和逻辑推理能力，甚至还要学会数据可视化技能，等等。在这里，或许我们可以用中世纪早期的英国小说《希尔德》（*Hild*）中的主人公来作类比。[35]希尔德天生聪慧，因擅长织造挂毯而为人熟知。说起她操作织机、飞速织毯的技巧，人们都形容她“眼睛一转就能创造出一个新花样”。对于她

的智慧，人们的评价亦复如是。这种建模思维（pattern-making mind）的本质特征是：能够创建联系，并看清互联网络。正是因为拥有这种能力，希尔德才得以游刃有余地周旋在种种地缘政治的阴谋中。

当然，如果让通才“单独行事”，那么他们够发挥的作用也相当有限。只有在与专家一起工作时，通才才能发挥出最大的作用：他们能够在翻译和交流的过程中为专家提供很大帮助；或者对专家的工作进行补充，而这种互补性非常重要。由此可见，在企业或组织中，为每个大型项目配备通才，或者流动性的通才小组，是很有必要的。通才可以为整个项目提供必要的背景知识和价值。必要的时候，企业还可以设立一个通才部门，对外服务，与专业公司合作，形成互补优势。

通才并不仅仅是T型人才。T型人才只是比一般的专业人士拥有更广的知识面；但作为通才，还必须有意识地将多个有实质性区别的领域联系在一起，尽管这样做有可能会失败。最优秀的通才，能够将如编年体般的历史细节与建模思维完美地结合起来。拥有建模思维能力的人，不会是那些只擅长抽象化和一般化的人，而是那些善于创建联系和类比的人。在高度复杂的系统中，对各部分相互关系和交互作用的洞察是非常重要的。因而，拥有建模思维的人，通过直觉和生物学思维相结合的方法，便能够理解，至少是部分理解这些系统。

我是国际赏云协会（The Cloud Appreciation Society）的成员。在我眼中，尽管万里无云的碧空也很美丽，但如果能形成一定的对比，或许会更令人激赏。也就是说，天空需要云。在我第一次告诉我太太，我是赏云协会成员的时候，她以为那是一个稀奇古怪的组织。于是，我给她看了很多“证据”：终身会员资格证、小徽章和官方证书等，但这反而让她觉得更加不可思议。然而，我真的认为，云既美丽又迷人。不过，在这个方面，我的理解并不比任何一个热情的业余爱好者深入多少。

未经专业学习的人，或许知道云的种类有很多，却不一定清楚它们之间的区别。你可能会说，有暴风雨来临之前的云、蓬松的云、色彩斑驳的云等，然而这是相当幼稚的分类方法，不能充分反映出云的复杂性和多样性。按照正式的分类和称谓，我们常见的云包括：积云、卷云、雨层云、积雨云等；还有一些特定的子类型，譬如碎云、陆架云、荚状云等。事实上，对云的科学分类是在近几个世纪里才发展起来的。正如科普读物出版家约翰·普塔克（John Ptak）所说：“云‘逃过’了古往今来最伟大的分类专家亚里士多德的法眼，后来又很少受到其他科学家的关注。[36]而且，20多个世纪以来，没有人对云进行过科学分类，直到……

1803年，英国制造化学家、气象学家卢克·霍华德（Luke Howard）公布了他对云的分类。”

发生在天空中的其他现象也激发出了许多美妙的气象术语，比如片状闪电、圣埃尔莫之火（St. Elmo's fire）、球状闪电，不一而足，它们大多都很怪异。就气象科学而言，在预测天气和解释气象方面，都已取得了很大进展；但是纵观历史，新奇天象层出不穷，这又为我们思考“大气如何影响天气”提供了新的思路。

云和其他大气现象，虽然不是人类所制造，但我们仍然可以从对待它们的方式中，学到一些对待复杂技术系统的思维方式。当我们不理解某个现象或某个系统或心生畏惧时，不应该避之若浼。有些事情即使暂时无从解释，但也应有一席之地；虽然处在尚待理解的范畴，但也可成为深入研究的楔子。

在面对某个完全无法理解的事物时，可以暂且先关注该系统中的细节，尝试着去理解整体中的某些特定部分。说到底，我们在研究系统时所采用的生物学方法，其实就是迭代和修补。也就是说，细节和意外情况不但能够帮助我们更加深入地理解系统，还能不断提高我们的洞察力。

在复杂难解的事物身上，自有神奇曼妙之处。那个闪耀着炫目光芒的技术网络复杂得令人难以置信，因为其各个组成部分之间拥有着紧密的联系和相互作用。许多时候，我们无法了解它的每一个部分，也无法完全理解它的整体性，不过，能对它有个不完美的把握，或许已经足够。我们可以与技术携手，谦卑前行。

OVERCOM PLICATED

Technology at the limits of comprehension

认识复杂系统的正确态度是：对于难以理解的事物，要努力克服我们的无知；一旦理解了某个事物，也不会认为它是理所当然的。谦卑之心，加上迭代的生物学思维，就是洞悉复杂世界的正确方式。

06 生物学思维是理解复杂世界的一把金钥匙

认识复杂系统的正确态度是：对于难以理解的事物，要努力克服我们的无知；一旦理解了某个事物，也不会认为它是理所当然的。谦卑之心，加上迭代的生物学思维，就是洞悉复杂世界的正确方式。

摩西·迈蒙尼德（Moses Maimonides）在哲学史上是一个独具特色的人物。他生活在12世纪，从小就接受了成为医生和拉比的训练，后来在埃及做了一名宫廷医生。他不仅编纂了犹太法律，还对当时的科学思想进行了融合与创新。他的崇高地位在犹太传统教义中得到了充分肯定："从摩西到摩西，从来没有出现过像摩西这样的人物。"迈蒙尼德还著有一本哲学教科书，名为《迷途指津》（*The Guide of the Perplexed*）。这是一部伟大的作品，将犹太教思想与亚里士多德的哲学理论融合在了一起。这个尝试很大胆，毕竟在那个时代，亚里士多德的哲学理论被公认为最先进的哲学思想。无论在犹太人当中，还是在非犹太人当中，《迷途指津》都拥有非常多的读者。

那么，我们可以从这位学者身上学到什么呢？在好奇心的驱使下，迈蒙尼德几乎研究了从营养学、天文学到政治学的所有学科。他认识到，尽管人类拥有非凡的智力，但是仅靠好奇心是无法理解世间万物的；至少，在涉及"无限"时，人类的思维能力是必定不够用的。尽管迈蒙尼德的观念在很多方面都领先于时代，但他还是受到了时代的限制。

他认为唯一可以理解世间万物的只有“上帝”。当然，在迈蒙尼德的字典里，“上帝”是比“天上的伟人”更抽象、更具哲学意义的概念；而“天上的伟人”则是人对“神”的共同认知。迈蒙尼德坚定地认为，人类的思维能力具有局限性：“……人的智力，毫无疑问，是有限的；人死去后，智力也就消失了。因此，人们都明白，世间必然存在着某些自己无法理解的事物。而且，他们还发现自己的灵魂并不渴望这种知识，因为他们已经意识到了这种知识的不可理解性，意识到了获取这种知识的途径是不存在的。”[1]

几百年后，人类的思维方式已经缓慢而稳定地实现了转变。而今，我们用“科学革命”这个术语来定义这种转变。随着现代科学的发展，科学家开创了天体物理学和化学等新领域，这让我们认识到某些局限性是可以克服的。当然，这种转变绝不是突然发生的，毕竟像牛顿这样的杰出人物，也将一生中大部分的时间和精力花在了对炼金术和中世纪神秘主义的追求上。无论如何，总的来说，科学观念还发生了明显的转变，人们开始认为，人类的大脑原则上可以理解它想要理解的一切。

这种信心渐渐渗透到了我们对待科学的态度当中，而到了当代，这种信心已经升华为一种必胜的信念：只要足够努力，随着时间的推移，我们定能理解一切。事实上，在探索宇宙和寻找答案的道路上，我们确实已经走了很久了。

在科学方面，我们再次向人类极限发起了冲击。现在看来，想要回答某些关于宇宙本质的问题，恐怕不得不去做一些“不可能”的事，并且要以超光速旅行。另外，对量子世界的研究似乎也正在接近极限。换句话说，我们正在探索的问题是目前还无法回答的。用生理学家、生物化学家、群体遗传学家J. B. S.霍尔丹（J. B. S. Haldane）的话来说：“现在，我所怀疑的是，宇宙不仅比我们想象的要怪异，而且也比我们能想象到的要怪异。”[2]虽然不应该放弃探索，但是答案可能会停留在我们的认知范畴内，而我们的认知能力正越来越多地受到限制。[3]

对于自己所创造的事物而言，情况也是如此。不难看到，我们对计算、运输、医疗设备，以及其他许多技术的理解都相当有限。所以，当我们继续构建无从理解的事物时，应该重新审视一下人类早期的思想。在那个时代，人们普遍认为，有些知识是任何人都无法获得的。

那么，在这种情况下，我们应该作何反应呢？我们在本书导论中提到过，随着技术变得越来越复杂，我们渐渐失去了理解它们的能力，在这种情况下，我们的反应会趋于两个极端：要么恐惧，要么崇敬。

一方面，在面对异常复杂的科技系统时，一些人会被它的强大和复杂压倒，并因未知而感到恐惧；另一方面，在面对科技所展现出的美丽与力量时，另一些人则会像对待宗教信仰一般表现出无比崇敬。视频游戏设计师兼作家伊恩·博格斯特（Ian　Bogost）甚至认为，用“上帝”这个词取代“算法”这个术语，一点也不会改变“技术”一词在流行语中的风评。[4]

虽然技术推动了社会发展，但它的构建过程并不完美。[5]技术是拼凑而成的系统，这就是说，技术是在时间的长河中，由许多碎片积累、层叠而成的系统。因此，尽管技术拥有令人振奋的力量，但它并不值得我们毫无保留地卑躬屈膝、赞叹不止。当然，我们也不必因技术的存在而感觉自身受到了威胁。

恐惧和崇敬都不是生产性的反应。这两种反应都阻断了我们提出质疑的机会，磨灭了理解技术系统的可能性。虽然在处理中高难度的事情时，我们能做到谨慎有加；但在面对完全无法理解的事情时，我们常常会受制于恐惧，并停下探究的步伐，而且多多少少会忘记自身责任，忘记我们应该尽力去理解能够理解的事情。对技术的崇拜也会导致同样的问题。如果我们对一种算法或技术的评价言过其实，无论是说它“美轮美奂”，还是说它“刻骨铭心”，我们都会因此而失去质疑的机会。

幸运的是，在上述两种极端反应之间，我们还有第三条路可以走，那就是谦卑。我们必须保持谦卑，而不应盲目崇拜；必须满怀好奇，而不应心怀恐惧。对此，计算机科学家迪科斯彻甚至说“卑微的程序员”[6]必须尊重“人类心智的内在局限性”。即使我们能够消除心智上的各种偏差，或者大幅度地增强智力，我们的生命也是有限的。面对这种终极的有限性，我们必须心存谦卑。谦卑既不是桀骜不驯，也不是委曲求全，而是承认自己的局限性，同时不会因这种局限性而陷入无为，也不会将这种局限性奉为神祇。以谦卑之心对待技术系统，有助于我们理解各种各样的系统结构，当然，这种理解仍在我们的认知范畴内。这种对待技术的谦卑态度与生物学思维十分合拍。虽然不断积累的知识让我们对系统有了更多更好的理解，但是这种迭代过程永远都是不完整的。当然，这没有什么关系。

《纽约时报》专栏作家戴维·布鲁克斯（David　Brooks）曾经明确指出：“智慧始于认识论中的谦虚。”[7]我们心怀谦卑，并对复杂系统的细节充满浓厚兴趣时，就能做到那些满怀恐惧或崇拜之心的人无法做到的事情。谦卑和兴趣能够帮助我们不断探寻系统背后的奥秘，即使永远无

法完全理解系统也没有关系。在许多情况下，部分理解已是我们能达到的最好的效果了。毕竟，聊胜于无。

恐惧和崇敬皆会诱使我们举手投降，放弃对技术系统的努力探索，但我们绝不能这样做，必须继续努力，尽力去了解这些系统。谦卑仅仅意味着我们需要接受以下观点：科学必胜论是错误的。我们永远无法实现完全或完美的理解。只要接受了这一点，我们的灵魂就不会再渴求这样的知识了，正如迈蒙尼德当年所说的那样。而后，我们会变得更加“通情达理”，不会再为无法理解的东西感到沮丧。当今时代，即使是软件专家也不得不承认，某些计算机错误属于形而上学领域。因此，是时候用谦卑之心来对待技术了。

事实上，在面对复杂的技术系统时，谦卑和包容是非常值得尊重的选择。研究生命基因组的科学家就是这方面的典范。在基因组中充斥着无数无“目的”的片段，它们是在奇异的吸积过程中形成的。在技术系统中也有类似的过程和片段，而科学家却认为这是“光荣的混乱”。[8]不难看出，用“光荣”来形容“混乱”，完全是生物学思维，满含谦卑与钦佩。技术系统是混乱的，是永远无法被完全理解的，但我们必须认识到，这就是“光荣的混乱”。这是一种非常有价值的乐观主义精神：繁杂的细节和不完美的理解都无法压倒我们，反而会令我们兴奋不已。也许它们永远不能为我们带来对整个系统全面且深刻的理解，但那不也很好吗？

约翰·盖尔（John Gall）是一位退休儿科医生，也是畅销书《系统圣经》（*The Systems Bible*）的作者。这本书原名为《一般系统滑稽论》（*General Systemantics*），于1975年首次出版，经过修订和扩充后的最新版第3版前段时间面市了。这本书对“如何处理复杂系统”做出了妙趣横生的探索，不过盖尔所指的“系统”具有更加宽泛的意义，涵盖了所有的社会系统和人类所构建的技术系统。书中金句连连，比如，“旧系统的幽灵会不断骚扰新系统”“系统总是会反击”。书里还提到了无意识定理（Unawareness Theorem）：“如果你压根不知道你已经出了问题，又怎么可能会寻求帮助？”[9]盖尔总结的定理以及对系统的分析都非常有见地且非常有趣。而且从这本书的原名出发，我们不难推测出，他的基本结论就是：系统很容易做出一些古怪的行为和出人意料的事情，比如对我们“反咬一口”等，而我们不可能根除这种行为。

透过那些金句，我们可以看到，盖尔提出的许多观点，与我们在本书中所阐述的思想类似。[10]例如，系统会吸积、会扩张、会发展，并最终“超越人类的评判能力”，做出意想不到的行为。不过在这里，我将只

针对其中一部分观点进行讨论。盖尔指出，从手头已有的东西开始构建，要比从头开始更加容易，因为从头来过所导致的问题可能比预期的要多很多。如果不得不从头开始创建一个全新的系统，那么我们要尽可能地控制系统规模。他认为，在减少系统故障方面，某种程度上我们还是可以想到一些办法的。

说到底，盖尔提出的金句和定理可以归结为一点：在面对很难从头开始设计或构建的系统时，务必要保持谦卑。[11]无论这些系统的起源或功能如何，它们最终都会变成笨拙的拼凑而成的系统。同时，在试图理解和控制它们的过程中，我们必须容忍许多混乱之物。只有在充分认识到“系统总会变得复杂”的时候，我们才能更好地构建系统，才能在复杂系统出现各种情况时，不至于被它们“反咬一口”，并且能够应对自如，甚至欢欣鼓舞。

在复杂的技术系统面前，保持谦卑将使我们获得良好的服务。这是一种充满智慧的谦卑，它的关键特征之一来自生物学思维的如下洞见：对于这些复杂系统的内部运行情况，窥一斑而难知全豹；或者说，我们只能在庞大的机器上打开一扇小门。但我们尽力了，这便已足够。

不要被表象迷惑

设计师唐纳德·诺曼（Donald Norman）正在将电脑中的文件备份到一台服务器上。他坐在旁边，观察着进度，想看看计算机一步步地都做了些什么。[12]一段时间后，他注意到备份进度已经来到了“样条网格化”（reticulating splines）阶段。这个专业术语看上去很高级，也显得很令人放心——计算机程序肯定真的知道它自己在做什么。但是诺曼没有感到放心，他想知道这到底是什么意思。经过一番研究后，他发现，正如《模拟城市2000》（*SimCity 2000*）游戏的拥趸们都知道的那样，这实际上只是一个“圈内笑话”。这款游戏被插入了一些无用的短语，只是看起来有意义罢了。自那以后，这种情况便陆续出现在了各种各样的游戏和其他软件中。

不妨回想一下，上次安装新软件的时候，你真的知道计算机内部发生了什么事情吗？你是否了解各种软件包被存储在硬盘驱动器上的哪些文件夹中？你知道软件安装程序根据计算机硬件和操作系统的不同性能，修改了计算机的哪些信息吗？

不太可能。你眼中只有一个长条状的东西，你只能通过观察它被逐

渐填满的过程，来了解安装的进度。它就是进度条。这种小型的界面创新程序，出自计算机科学家布拉德·A.迈尔斯（Brad A. Myers）之手。在读研究生时，迈尔斯开发出了这个程序，并将其称为“进度完成指示器”。[13]进度条给不透明的过程提供了一个小窗口，并以此来抚慰用户。那么，进度条是不是完全准确的？很可能不是。在某些情况下，进度条几乎完全脱离了底层的运行过程。[14]但是在大多数情况下，进度条和其他讨巧的设计，例如描述软件安装过程中所发生之事的文本框一样，可以让人们对一个庞大且复杂的过程投去“放心的一瞥”。

我们在构建用户界面时会越来越多地有意地将复杂性抽象掉，或者，至少让部分复杂性与用户隔离。这种做法使复杂性科学和用户界面设计学这两个领域产生了交集。无论是在电脑中，还是在汽车中，抑或是在电器中，技术系统都会在用户与程序运行之间设置一道屏障。许多人都会选择使用用户界面相当友好的特波税务软件（TurboTax），而不会选择与越来越复杂的税法进行面对面的“博斗”。[15]然而，在这个软件的背后，其实隐藏着一整套非常复杂的法律法规体系，它们是通过大量if语句和处理例外情况的计算机代码来呈现的。

凭借直觉，我们能察觉到系统内部到底发生着什么，尽管这种直觉有时候不甚准确。因此，只要想办法保证这种直觉不消失，就可以消除用户因未知而产生的不安。

我家的第一台电脑是康懋达公司生产的VIC-20。在其代言人、《星际迷航》（*Star Trek*）的主演威廉·夏特纳（William Shatner）口中，它是“20世纪80年代的奇迹电脑”。我拥有许多有关这款古董级电脑的美好回忆。我曾经用它玩过不少游戏，当时所用的存储设备还是最原始的盒式磁带机。我和我的兄弟玩过很多次的一个游戏是《吃豆人》（*Pac-Man*），不过，在那款游戏中，“吃豆人”的形象不是一个长着大嘴的黄色馅饼，而是一辆赛车。

我们还玩了不少需要手动输入代码来运行的游戏。在那个年代，虽然你也可以购买现成的VIC-20游戏软件，比如赛车游戏，但是人们玩电脑游戏的主要方式是将杂志上发布的计算机代码输入到电脑里。你想玩一个有趣的滑雪游戏吗？买一本杂志，将上面的计算机代码输入到电脑中，然后自己来运行这个程序，并不需要购买游戏软件。当时，这样的游戏很常见。今非昔比，如今，几乎每个游戏程序都有几十万行代码，可能整本杂志都登不完。

自己输入代码，可以让我们更近距离地接触计算机。我看到了软件

中的bug是如何产生的。记忆中，我输入的滑雪游戏的程序代码导致屏幕一侧出现了很多图形乱码，我费了很大心思才把这个错误纠正过来。我还看到了计算机程序的逻辑和内在结构。今天的计算机程序是以成品的形式交付给用户的，相比之下，往日的计算机程序则具有清晰的结构。

20世纪80年代末，我的家人放弃了那台康懋达公司生产的VIC-20，转而使用苹果公司的产品。自那以后，我一直在使用苹果电脑。我们家的第一台苹果电脑，在年幼的我们看来是一个令人难以置信的产品。无论是操作便利的鼠标，还是大量的游戏，都深深吸引着我。我记得其中有一款游戏叫《宇宙渗透》（*Cosmic Osmo*），它拥有内涵丰富的游戏场景，令人顿生身临其境之感，而且只需点击鼠标即可尽情探索。早期的苹果电脑也会“说话”，能将文本转换为单调的“非人类”语言，这让我们全家都感到无比兴奋。1984年，史蒂夫·乔布斯在苹果电脑新品发布会上的演讲感人至深，给我留下了难以磨灭的印象。然而，在拥抱苹果电脑这个奇迹的过程中，我和我的家人也失去了一些东西。我们离计算机越来越远了，而这种趋势在今天仍在继续：在用惯了顺手的iPad之后，我们甚至忘记了文件是怎样存储的。

不过，我还是给家里的苹果电脑安装了HyperCard。HyperCard是编程语言和开发环境的奇妙结合。你可以创建出虚拟卡，然后将它们拼接在一起，并添加表示特定功能的按钮和图标；你还可以制作出有趣的动画和酷炫的声效，然后把它们关联到其他卡上。或许你曾经玩过经典游戏《神秘岛》（*Myst*），它就是用HyperCard开发出来的。HyperCard就像一系列被保存在个人电脑上的原型网页，可以帮你做想做的任何事情。在当时，对于我这个刚开始探索计算机世界的孩子来说，这个可视化创作空间可谓是通向未来的完美窗口。

我用HyperCard创建了一个相当原始的密码生成器，它可以生成能用作密码的随机字符串，而且还带有一个可以让随机密码更易读写和长期记忆的选项。这个当然是比较简单的，但是在当时的我看来，它绝对领先于时代。

计算机游戏设计师哈伊姆·金戈尔德（Chaim Gingold）将HyperCard这类入门级软件工具称为“魔术蜡笔”。在童话《阿罗有支彩色铅笔》（*Harold and the Purple Crayon*）中，主人公哈罗德有支神奇的彩笔，用它画出什么，现实世界中就会出现什么。在金戈尔德看来，HyperCard就是这样的神奇工具，可以让“非程序员也能参与数字媒体的

创作，并构建出动态的事物”。[16]人们普遍认为，苹果电脑已经“取消”了杂乱的代码，隐藏了计算机内部的运行状态，但即使真是如此，用户依然可以通过HyperCard造访复杂的数字领域。就我自己来说，HyperCard为我提供了一个可以窥探数字世界深处的舒适窗口，并为我提供了一些系统内部运行的线索。

能与我们产生交互的所有复杂系统都是有层级的。在技术系统中，我们通过深思熟虑的抽象化方法构建了这种分层结构；在自然界中，分层结构则建立在规模和进化的基础上。生物学中的层级从生化酶开始，到线粒体、细胞、器官，再到整个生命体，而后是整个生态系统。每个层级都可以提供不同层面的信息。当从一个层级的抽象化，向上扩展到另一个层级的抽象化时，我们或许会失去对细粒度（fine-grained）的理解和控制，但同时也能更好地理解更高级别的体系。在计算机软件中，我们的关注点可以从比特与字节转移到汇编语言，再到更高级别的计算机代码，最后到日常的用户界面，也就是我们点击、拖动的东西，例如对浏览器的使用。每一个更高的层级会都为我们带来更多的功能，但同时也会让我们更加远离底层逻辑。

当然，现在我们已经很清楚了，人类是不可能完全理解底层逻辑的。但是，我们至少应该窥探一下机罩底下的世界。如果将日常使用的平板电脑和手机视为会变魔术的玻璃和金属抛光板，完全不去理会面板下面或数字键下面到底发生了什么，那么我们肯定会失去一些机会。事实上，这确实是个隐患：如果所使用的系统是完全自动化的，[17]那么当它出现问题时，我们几乎无能为力。我们被屏蔽了，无法了解技术系统的内部运行，这种状态被称为“隐藏的电子复杂性”。换句话说，我们所用的设备存在着惊人的复杂性，但这种复杂性却被完全隐藏了起来。[18]

在20世纪60年代的电话通讯系统中存在着这样一个设计：当系统检测到故障时，就会将用户拨号关联到错误的号码上。[19]这个设计的目的是将技术系统故障定向为人为错误，以逃避人们对技术系统的苛责，用户会认为是自己拨错了号码。因为不了解内情，所以那些“拨错了号码的人”对电话通讯系统的权威性和神秘性有着自己的理解；一旦让他们知道了电话通讯系统是庞杂的、不完美的“人造物”，他们的理解就会变得不一样。

是的，我们必须打开机罩或通过窗口往里面瞥上一眼，看看那里发生了什么，尽管这种“管中窥豹”的做法通常无法做到“知全豹”。当技术的复杂性超出了人们的理解能力时，这样的“一瞥”，无论是对专家而

言，还是对普通用户来说，都至关重要。仅仅让一个人，或一小部分人看到底层逻辑，并认识到系统和人类的局限性，是远远不够的。人人都应该关注，如何实现这“一瞥”。倘若做不到，我们在系统面前便会失去谦卑之心和敏锐直觉，从而转向崇敬或恐惧。关注技术底层，不仅有趣且有教育意义，而且还有助于抵制对待技术的“不正之风”。在当今这个纠缠时代中，我们越来越需要这种抵御能力了。

不过，如果很难瞥见机罩下的东西，我们又该怎么办呢？如果一个系统复杂到难以构建窗口，或是所提供的线索太少，我们又该如何是好？这里可以考虑另一种方法：构建模型。通过构建模型，我们可以推测出一个复杂的技术系统是如何运行的。

例如，虽然我们无法真的控制气象，也无法理解它所有的非线性细节，但却可以相当好地预测它、适应它，甚至为坏天气的到来做好准备。气象模型非常复杂，尽管它每个部分的设计都不是很难。我们不仅希望利用气象模型来指导日常活动，比如何时清理衣橱，何时出远门等，还希望对气象模型有所理解，哪怕是不完全的理解。当然，在面对突如其来的暴风雪或洪水时，我们还希望能靠它避险。

就像构建气象模型一样，我们也可以构建技术系统模型，哪怕有所简化也没关系。通过对感兴趣的系统进行模型构建，并测试其极限，调整其参数，我们便可以获得对复杂系统的强大洞察力，而用不着完全理解它。而且，构建模型本身也是一项值得培养的技能。

电脑游戏《模拟城市》就是这样的一个模型。[20]通过观察城市的运转模式，用户的洞察力得到了锻炼。在《模拟城市》出现之前，我一直不相信，没有城市规划或土木工程等专业背景的普通人，对城市的运转模式也能有清晰的理解。当然，我们无法扭转城市发展的进程，也无法进行反事实的实验。在《模拟城市》出现之后，我们或许仍然无法完全理解现实中城市的复杂性，但是一般来说，玩过这个游戏的人对城市的运转模式确实有了更深刻的理解。我们不仅需要更好的与技术世界有关的模型或游戏，还要教会学生如何“玩”一些系统，测试它们的极限，探析它们的工作原理，至少在某种程度上去探析。这种“玩”，其实是模拟技术故障，并观察系统反应，可以让我们更舒适地与庞大且笨拙的系统相处，并更好地理解这个日益复杂的技术世界。

我们还需要一些类似于电视上的气象学家的，能够对系统内部事态作出说明的解释者。经济学家泰勒·考恩在其《平均时代的终结》一书的结尾部分曾预言，这种解释者即将登场。[21]在他看来，他们将会“磨炼

自己寻找、吸收和评估这方面信息的技能……他们将会成为机器网络世界的真相阐释者……至少，在未来相当长的一段时期内，只有他们能够清楚地认识到这个世界所发生的事情”。这些解释者对模型构建的评价可能会相当高，因为那可以给我们带来一些关于复杂系统的灵感。[22]随着时间的推移，技术将变得越来越难以渗透，为此，中学和大学教育需要更多地引入模型构建和简化直觉，并把它们结合起来，作为必要的复杂性科学方法传授给学生，让每个人都有机会成为系统的解释者。

简化直觉是由复杂性科学家发展起来的一般模型，它为启发式方法奠定了基础。启发式方法可以使我们得以理解系统，同时又不会被其复杂性压倒，并能与“包容性的将就”相结合，这也是约翰·盖尔的“一般系统滑稽论”所采用的视角。这些用来窥探大规模复杂系统的小窗口，体现了我们对待技术的谦卑之心。

另外，在面对复杂的技术系统时，我们还可以采用另一种更加积极的思维方式。

以欣慰感看待不理解的事物

在意第绪语（Yiddish）中有一个非常有意思的词naches。这个词的含义是为他人感到骄傲或喜悦，也就是所谓的欣慰感。在意第绪语中，人们经常说，要学会shep naches，意为要学会从亲近之人，尤其是自己孩子的成就中捕获替代性的自豪感。这是最纯粹的愉悦感，也是你在酒吧、婚礼和毕业典礼上经常听人说起的一种快乐感受。

移民者一般都有一个执念：让下一代生活得比自己好，让第三代、第四代繁荣兴旺。对于父母来说，后代“必须”比自己更聪明、更成功。我们之所以会因下一代的成功倍感欣慰，原因之一便是我们受到了这个执念的驱使。那么，我们为什么不可以到技术成就中去捕获这种欣慰感呢？

我们或许不了解机器或系统正在做什么，在某些情况下，它们甚至可能是由多种技术构成的，但无论如何，它们都是人类智慧的“儿子”，是我们智力的“后代”，因而技术成就也可以令我们倍感欣慰。数千年以来，人类一直非常珍视这种情感，它给我们带来了巨大的快乐。现在，我们只需要将为人父母的自豪感转移到技术领域即可。

那么，这种欣慰感对于技术而言究竟意味着什么？最基本的，我们

这些机器的创造者可以从技术“后代”的“成就”和“能力”中获得快乐。以计算机程序为例，除了其他成就之外，计算机程序现在已经能够创作出复杂的绘画作品和音乐作品了。加利福尼亚大学圣克鲁兹分校的作曲家戴维·科普（David Cope）所开发的一个软件能够按照给定作曲家的风格创作出新颖的音乐，而且这些音乐作品听起来非常不错。[23]比如，由这个软件创作的美国作曲家斯科特·乔普林（Scott Joplin）风格的作品，听起来就像是乔普林本人创作的一样。尽管科普并没有直接参与创作，但他仍然会为这个程序的“成就”感到自豪。他一手打造的计算机程序所创作的作品为他带来了很大的快乐。[24]同样，超级电脑沃森（Watson）的创造者、IBM的程序员们也会为沃森在国际象棋比赛中战胜了人类棋手而欣喜若狂。

我们可以将这种欣慰感延伸到更广阔的领域中。许多人都有自己支持的球队，并会为它的胜利而倍感自豪，尽管其实他们与球员们并无交集。当自己国家的运动员在奥运会上夺冠时，当自己国家的公民因重大发现而获得权威奖项时，我们都会感到兴奋无比。同样的道理，人类也可以为机器所取得的“成就”而感到非常欣慰，毕竟它们是由人类创造的，即使它们的成就不是我们的个人成就，即使我们无法完全理解它们。事实的确如此。从苹果手机到互联网，技术进步让许多人倍感欣慰，并心生感恩，尽管他们不一定了解这些技术的工作原理。

即使我们所构建的机器的复杂性超出了人们所能理解的范畴，我们也无须感到不安或失望。当我们的孩子创造出了令人惊讶的作品，而我们却无法理解时，我们是不会感到绝望或担心的；相反，我们会为他们的成功感到高兴，甚至会心生感恩。对于技术系统的“作品”，我们完全可以给予类似的回应。

“欣慰感”是一个框架，能够帮助我们认识到，我们正循着祖先数千年来的足迹前进着。[25]在前行的道路上，越来越少有人能够理解“人造世界”中最复杂的部分了。较之以往，最近的趋势也算不上新鲜或不同。

通过“欣慰感”这个镜头折射出来的谦卑之心，实为一种乐观主义精神，它让我们对自己所构建的这个“不可理喻”的世界心怀希望。除此之外，谦卑还能让我们在不断的追求中始终保持平衡：既能把握所构建事物的所有维度，又能看清自身的局限性。当然，对于专家来说，这是一种使命。

而且，谦卑还能帮助我们看清神秘与奇迹之间的区别。

谦卑之心 + 迭代的生物学思维

《奇迹世界》（*The World of Wonders*）是一本出版于19世纪下半叶的书，书中汇集了许多新奇的想法、迷人的事物和离奇的事件。[26]从创作风格和写作手法看来，作者显然是一个热爱生活、享受生活，并充满好奇心的人，并不断地将人类进步与自然奇观交织在一起。这本书的内容涉及“历史和哲学中的一切奇思妙想、科学奇迹、眼镜商的玻璃，以及化学家通过努力所揭示的动物生命奇迹”等。当然，那些都是维多利亚时代的奇迹。这本书还展示了人类在周遭世界中寻找奇迹的多种方式。但是，奇迹世界并不是无法解释的世界，我们的每一次进步、每一项技术，都令人着迷并充满了力量。除此之外，最重要的是，它们皆能被理解，一点儿也不神秘。

对于复杂的技术系统中的神秘性，人们通常会有两种潜在的极端反应。第一种反应是有意且极力地削弱它的重要性，认为系统中根本不存在任何神秘的东西。对于系统的运行方式，这些人会编织出一些简单的故事。这些故事也许很吸引人，但都严重削弱了系统的复杂性。[27]他们认为，自己对正在使用的技术了如指掌，任何问题都是能解决的小问题。许多大型的科技公司都是这种风格，他们把所有意想不到的系统行为统统归结为小问题，还认为系统自身已经在处理这些问题了。

另一种极端反应则来自那些醉心于寻找神秘和未知事物的人。这些人并不在乎这种事物是否真的存在。他们通常都是技术的门外汉，总是把设备或系统的内部运行神秘化，并以此为荣；他们对苹果手机或电网的工作方式赞不绝口，并称其为魔术。这些人会问自己的父亲“电灯和真空吸尘器的工作原理是什么”，并满足于“简直就是魔术”之类的答案。不过，大多数人的反应都处于上述两种极端反应之间。他们可能会承认，在系统中存在一定程度的神秘性，但同时又希望自己可以理解这个世界。

当我们把奇迹和赞美，以及对神秘世界的感知都考虑进去时，事情就变得更加混乱了，我们甚至可能会“喜欢上”某种现象的复杂的美感。但是，从《奇迹世界》这本书中可以清楚地看到，奇迹不一定要以牺牲可理解性为代价，反之亦然。事实上，许多人都发现，在面对一个非常庞大且复杂的系统时，自己理解得越透彻，应对得越自如，内心就会越愉悦。

例如，唐·诺曼曾经写道，如果能真正理解棒球比赛中的内野高飞球

[28]规则，那将是一件多么令人开心的事情啊。棒球比赛的规则非常复杂，其规则手册多达200页，其中关于内野高飞球的规则还包含了大量的例外情况和补充说明：“在两人出局以前，当第一、二垒有跑垒员，或第一、二、三垒都有跑垒员时，击球员击出界内高飞球，且不含平飞球，也不含试图触击，同时内野手在正常守备条件下可以接住该球时，这个球即为内野高飞球。”棒球迷们花了很多的时间和心思去理解这个复杂的规则，有的人甚至觉得这里面隐藏着奇迹般的东西。

那么，我们在面对自己所构建的大型系统时，又该如何平衡其间的神秘和奇迹呢？特别是对于那些无论我们多么努力都无法完全理解的系统。

我们绝不能屈服于“将这两种感觉融合到一起”的想法，那会将神秘与奇迹混淆起来。[29]一些人认为，未能解释的东西定会激励我们继续前行。他们把神秘的东西错误地等同于神奇的东西，并据此认为，未能理解的东西也定是令人惊叹的。同样，我们也必须抵制“一旦被完全理解，就不再是奇迹”的想法。技术变革常常会引发这种思想。技术的发展日新月异，以致我们忘记了，我们对宇宙的理解以及理解宇宙的能力已经变得何其令人震惊；也忘记了其实直到最近我们才拥有了在屏幕上呈现三维图像，在地球上不间断通信，以及通过点击鼠标观看数十年前的电视广告的能力。不仅如此，那些让“不可能成为可能”的基础知识，也没能激发起人们心中的敬仰之意和惊叹之情。到最后，我们能感受到的只有“一个已经被看透、被控制的世界所留下的悲伤的惰性”。[30]这就是哲学家休伯特·德莱弗斯（Hubert Dreyfus）和肖恩·多兰斯·凯利（Sean Dorrance Kelly）所描述的那种风险：随着科技力量的崛起，人们心目中的奇迹将出现得越来越少，以致一些人只能成天幻想“迷人的过去”。

我们必须努力维持这样两种相互对立的状态：神秘而无奇迹、神奇而不神秘。第一种状态要求我们努力克服自己的无知，[31]绝不能沉迷其中；第二种状态意味着，一旦理解了某个事物，我们就不会认为它是理所当然的了。

在我看来，从本质上说，上述两种状态的最终诉求都是建立科学的思维模式，而科学的思维模式是我们学习新事物和解决难题的必备能力和先决条件。如果我们对自己无意中创造出来的神秘事物抱有太多好奇的话，那么原本想要“消除神秘事物、厘清纠缠时代”的想法就会被削弱。

我们永远无法完全理解那些拼凑而成的复杂系统。我们始终会面对

一些难解之谜，但是没关系，只要不害怕、不沉迷，就可以找到最好的方式——谦卑，哪怕携有一丝崇敬也无所谓，毕竟那是我们自己所构建的系统。谦卑、好奇，以及欣慰感，或许就是我们能做到的最好的状态了。我们必须坚持运用生物学思维，即使尚未实现对系统的充分理解。如果失败了，我们可以再谦卑一点，但绝不能气馁；我们定会找到让直觉深入技术系统底层，窥其堂奥的途径。

当丰田公司生产的汽车出现意外加速事故并导致死亡时，专家们无法准确地判断出故障原因，但他们给出了一个建议：通知相应型号丰田汽车的驾驶员，如果汽车出现自动加速并失去控制的情况，最快捷的解决方案就是立即将汽车挂到空挡上。[32]当然，作为一个消费者，你已经将家人和朋友，以及自己的生命“托付”给了一辆汽车，而这台强大的机器却可能以如此不可预知的、可怕的方式行事，那么恐怕你并不想听到这样一个建议。这就像是你摔断了骨头，庸医却建议你贴上一个创可贴。

但是，如果我们能明白，罕见的故障和无法解释的错误都是复杂性的衍生物；如果我们能知道，即使是最警觉的设计师也无法预测全部故障和错误，那么在遇到这种状况时，我们就能从容不迫，而不会手足无措了。这也就是说，如果消费者能认识到，这种混乱的、难以理解的系统是现实新世界的一部分，[33]那么就不会再认为“立即换成空挡”是面对故障时最糟糕的选择了。[34]事实上，在当今时代，设计师和驾驶员已经成了探索纠缠时代的搭档，至少在以下这个非常重要的方面，他们并不像看起来那样毫无关联：谁都不可能完全预测复杂系统的所有行为。许多人都认识到了这一点，只是采用了不那么危险的处理方式，比如说，当计算机出现故障时，有些人会在沮丧中下意识地敲打敲打机身，或者重启计算机，希望系统能够自行解决问题。

进入20世纪后，我们迎来了无数的限制性定理（limitative theorems）。这些定理为我们划定了可知和可理解的知识边界。在数学领域，库尔特·哥德尔（Kurt Gödel）证明了，在某个给定的数学系统中，总是存在永远不能被证明为真或为假的陈述；在计算机科学领域，艾伦·图灵证明了，任何机器所能做到的事情都存在极限，无论某个特定的算法可能会发展到多么强大的程度。显然，这些领域并没有因此而消亡，甚至算不上遭遇了巨大挫折。尽管受到了限制，但它们在许多方向上都结出了累累硕果，其发展远远超出了那些为它们划定边界之人的想象。

在构建和使用复杂的技术系统时，我们试图了解它们如何运行以及如何失败，但是这种了解确实存在着极限。不过，这并不意味着我们要停止创造；恰恰相反，它只是意味着，当我们继续构建这类系统时，要明白它们还会不断成长，会变得更加奇异、更加复杂。因此，我们必须重新调整期望值。

从现在起，承认自己无法完全理解这些系统吧。这种做法将会改变我们理解和应对系统的方式。如果你参与了构建，那么在组装系统时，你就会意识到自己的理解能力是有限的；如果你与系统进行了交互，那么你会认识到繁杂和意外是常态，而不是例外；如果你试图彻底消除这种常态，那么你不但会失败，还会让事情变得更糟。

科幻小说作家威廉·吉布森（William Gibson）经常提到“不可思议的现在”（unthinkable present）这个短语，这也是他笔下众多故事的背景。[35]无疑，这恰是我们所处的这个纠缠时代的内核。但是说到底，这些系统并不是完全无法被理解的，其性质和特征也不会永远不可言说。这个故事并没有以“对不可理解之物的恐惧”或“对不可思议之物的好奇”而结束。谦卑之心，再加上迭代的生物学思维，就是洞悉复杂世界的正确方式，会指导我们与复杂系统进行交互。但是，无论我们所持观点的细节如何，我们都无须屈服于路德维希·维特根斯坦的逻辑：“凡是不可言说之事，我们必须保持沉默。”

事实上，对于那些无法被完全理解、无法被完美处理的事物，我们还有很多话可以说！人类，将在这个纠缠时代继续生存下去，并发展壮大。

注释

引言　认识复杂系统

1. 有人猜测，《华尔街日报》的网站之所以会崩溃，是因为有太多用户为了阅读关于“纳斯达克停止交易”这则新闻而在同一时间段内登陆。请参阅：Jose Pagliery, “Tech Fail! Explaining Today's 3 Big Computer Errors,” *CNN* Money, July 8, 2015。关于此事件及其影响的进一步分析，请参阅：Zeynep Tufekci, “Why the Great Glitch of July 8th Should Scare You,” *The Message*, July 8, 2015。

2. Andrea Chang and Tracey Lien, “Outages at NYSE, United Airlines, WSJ. com Expose Digital Vulnerabilities,” July 8, 2015.

3. Thomas Homer-Dixon, *The Ingenuity Gap: Facing the Economic, Environmental, and Other Challenges of an Increasingly Complex and Unpredictable Future* (New York: Alfred A. Knopf, 2000; repr. Vintage, 2002), 171.

4. Edsger W. Dijkstra, “On the Cruelty of Really Teaching Computing Science,” E. W. Dijkstra Archive: The manuscripts of Edsger W. Dijkstra, 1930—2002, document no. EWD1036, December 1988 . 艾兹格·迪科斯彻在1972年受美国计算机协会（ACM）之邀举办了图灵讲座（Turing Lecture），并阐述了类似观点。请参阅：E. W. Dijkstra, “The Humble Programmer,” *Communications of the ACM* 15, no.10 (1972): 859–86。尽管不是每一项技术都具有可计算性，但迪科斯彻的确深刻地指出了当今这个技术时代的生活模式。

5. Lee Billings, “Brave New Epoch,” *Nautilus* 009: January 30, 2014.

6. Neil Johnson et al., “Abrupt Rise of New Machine Ecology Beyond Human Response Time,” *Scientific Reports* 3: 2627, September 11, 2013.

7. 原文为：回到2008年，那是布拉德·胜山（Brad Katsuyama）第一次意识到股票市场已经变成了“黑箱”，普通人不可能理解它的内在工作机制。请参阅：Michael Lewis, “The Wolf Hunters of Wall Street,” *The*

New York Times Magazine, March 31, 2014。

8. Berkeley J. Dietvorst et al., “Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err,” *Journal of Experimental Psychology*: General 144, no.1(2015), 114–26.

9. 关于这种崇敬心理的实例有：一位记者写下了自己在参观谷歌数据中心时的感受；请参阅：Stephen Levy, “Google Throws Open Doors to Its Top-Secret Data Center,” *Wired*, October 17, 2012。另一位记者则描述了自己对Facebook新闻推送算法的崇敬之心，他是这样说的：“新闻推送算法在评论界掀起了轩然大波，它似乎有了自己的思想，就像是有智慧的北欧古符文被解除了封印，开始追逐一些超出人类理解能力的目标。”请参阅：Will Oremus, “Who Controls Your Facebook Feed,” *Slate*, January 3, 2016。

10. 关于“对算法的崇拜”的进一步讨论，请参阅：Ian Bogost, “The Cathedral of Computation,” *The Atlantic*, January 15, 2015。

01 欢迎来到这个纠缠的时代

1. James Gleick, “Richard Feynman Dead at 69; Leading Theoretical Physicist,”*The New York Times*, February 17.

2. 关于丰田汽车意外加速事故的更多信息，请参阅：Ken Bensinger and Jerry Hirsch, “Jury Hits Toyota with $3-million Verdict in Sudden Acceleration Death Case,” *Los Angeles Times*, October 24, 2013; Ralph Vartabedian and Ken Bensinger, “ Runaway Toyota Cases Ignored,” *Los Angeles Times*, November 8, 2009; Margaret Cronin Fisk, “Toyota Settles Oklahoma Acceleration Case After Verdict,” *Bloomberg Business*, October 25, 2013; Associated Press, “Jury Finds Toyota Liable in Fatal Wreck in Oklahoma,” *New York Times*, October 25, 2013。

3. 2014年10月3日，库普曼在他的博客上发表了一个题为“丰田汽车意外加速事故与软件安全案例研究”的演讲。也请参阅：the report by Michael Barr of the Barr Group on *Bookout v. Toyota*。

4. 在内在固有复杂性与偶发复杂性之间，存在着一条界线。前者是系统正常运行所需的复杂性，包括为例外情况和特殊情况提供预案的各种机制；后者则更接近于本书所说的“过度复杂”，往往出现在系统不按计划

超速"生长"的时候。

5. 关于阿丽亚娜5型火箭的事故详情，请参阅：Homer-Dixon, *The Ingenuity Gap*。霍默 - 狄克逊的部分叙事基于James Gleick, "A Bug and a Crash: Sometimes a Bug Is More Than a Nuisance," 1996(which originally appeared in *The New York Times Magazine*, December 1996)。关于这类系统突然崩溃的直接原因与深层原因的讨论，请参阅：Chris Clearfield and James Owen Weatherall, "Why the Flash Crash Really Matters," *Nautilus* 023, April 23, 2015。

6. Clearfield and Weatherall, "Why the Flash Crash Really Matters."

7. 从根本上说，这类事故皆可归因于内在复杂性，也就是存在于大型系统内部的，随时间推移而不断进化的复杂性，而不能归因于某个特定的外生冲击。

8. 我们必须认识到，"挑战者号"航天飞机的失事原因比坊间所传要复杂得多。例如，有种说法是，在发射之前，某些工程师已经预感到这架航天飞机可能会发生"灾难性事故"；尽管他们无法说出准确原因，但是他们的确反对过这次发射。当时，有工程师指出"温度是一个因果性的因素"，但并不确定。请参阅：Wade Robison et al., "Representation and Misrepresentation: Tufte and the Morton Thiokol Engineers on the Challenger,"*Science and Engineering Ethics* 8, no.1(2002): 59–81。在这个事故的"口述"材料中也有一些线索可以表明，某些工程师似乎知道失事原因；理查德·费曼也得到了这方面的材料，并在听证会上做出了强调。请参阅：Margaret Lazarus Dean, "An Oral History of the Space Shuttle Challenger Disaster," *Popular Mechanics*, February, 2016。

9. Philip Ball, "Science Fictions," *Aeon*, October 29, 2012.

10. Ian Beacock, "Humanist among Machines," *Aeon*, June 25, 2015.

11. Max Weber, "Science as a Vocation,"in *From Max Weber: Essays in Sociology*, trans. and ed. H. H. Gerth and C. Wright Mills, 129–56(New York: Oxford University Press, 1946, repr.1958) .

12. 人类一手创造之物，令人类困惑不解。类似的观点，法国诗人保罗·瓦莱里（Paul Valéry）早已"预言"："因此，所有的问题归根结底是在问：人类心智能不能掌握人类心智的创造物？"转引自：Langdon Winner, *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought* (Cambridge, MA: The MIT Press, 1977), 13。

13. 这只是两个术语之间可能存在的若干区别之一。

14. 感谢亚伦·克劳塞特。在我们对本书进行第一次讨论的时候，他为我提供了这个水上浮标的例子。

15. Kevin Kelly, “Cities are technological artifacts, the largest technology we make.” *What Technology Wants* (New York: Viking, 2010), 81.

16. David McCandless, “Codebases: Millions of Lines of Code,” infographic, v.0.9, *Information Is Beautiful*, September 24, 2015 . 假设一本百科全书有3万页，每一页可以放下1 000行代码，那么据推算，苹果操作系统需要好几本百科全书才能登载完。

17. 有信号灯的交叉路口数量是根据美国交通部联邦公路管理局2015年10月20日的数据做出的估计。

18. Gideon Lewis-Kraus, “The Fasinatng...Frustrating...Fascinating History of Autocorrect,” *Wired*, July 22, 2014.

19. Laura Saunders, “Paper Trail,” sidebar to article “Don't Make These Tax Mistakes: Fifteen Common Tax-Filing Errors That Can Cost You Dearly,” *The Wall Street Journal*, January 31, 2014.17.

20. 对复杂系统不同层级的理解，请参阅：Stephen Jay Kline, *Conceptual Foundations for Multidisciplinary Thinking* (Stanford, CA: Stanford University Press, 1995), 268。

21. Erik Sandberg-Diment, “A Computer Comes in from the Cold,” *The New York Times*, April 21, 1987.

22. Gerard J. Holzmann, “Code Inflation,” *IEEE Software* (March/ April 2015): 10–13.

23. Philip K. Howard, “Fixing Broken Government: Put Humans in Charge,”*The Atlantic*, September 22, 2014.

24. Quinn Norton, “Everything is Broken,” *The Message*, May 20, 2014.

25. Winner, *Autonomous Technology*, 290–91.

26. Danny Hillis, “The Age of Digital Entanglement,” *Scientific American*, September 2010, 93.

27. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford,

UK: Oxf ord University Press, 2014), 17．不过，此次闪电崩盘的原因至今不明。

28. César Hidalgo, *Why Information Grows: The Evolution of Order, from Atomsto Economies* (New York: Basic Books, 2015).

02 复杂系统形成的4个原因

1. Barry M. Leiner et al., “Brief History of the Internet,” Internet Society, October 15, 2012.

2. Randall Munrose, “DNA,” *xkcd*, November 18, 2015.

3. Chris Kirk, “Battling My Daemons: My Email Made Me Miserable. So I Decided to Build My Own Email App from Scratch,” *Slate*, February 25, 2015.

4. 计算机代码与法律代码、道德代码等其他代码所具有的相似性，还有很多可以讨论的，但是最好将它们视为不同的系统代码。

5. 请参阅：《美利坚合众国宪法》，第一条，第八款：“To Establish Post Offices and Post Roads”; “Title 39-Postal Service,” *Code of Federal Regulations* (annual edition),revised July 1, 2003。

6. Michael J. Bommarito II and Daniel M. Katz, “A Mathematical Approach to the Study of the United States Code,” *Physica A: Statistical Mechanics and its Applications* 389, no.19(2010): 4195–200.

7. “Wright 1903 Flyer,” NASA Glenn Research Center.

8. “747 Fun Facts,” Boeing Commercial Airplanes.

9. Kelly, *What Technology Wants*, 279.

10. Kelly, *What Technology Wants*, 279; David McCandless, “Codebases: Millions of Lines of Code,” infographic, v.0.9, *Information Is Beautiful*, September 24, 2015.

11. McCandless, “Codebases.”

12. M. D. Fagen, ed., *A History of Engineering and Science in the Bell System: The Early Years(1875—1925)*, Technical Publication Department, Bell Laboratories, 1975.

13. 美国联邦航空管理局担心"千年虫"问题的说法源自：Homer-Dixon, *The Ingenuity Gap*。

14. Matthew L. Wald, "Warning Issued on Air Traffic Computers," *The New York Times*, January 13, 1998．按照霍默－狄克逊的说法，他们雇用了一位退休程序员去帮助解决这个问题。

15. "IRS Trudges On with Aging Computers," *CNET News*, April 12, 2007；John Bodoh, "Tech Timebomb: The IRS Living in the 1960s," *Washington Examiner*, December 17, 2014.

16. "The Shuttle: NASA's IT Legacy," *Information Age*, July 18, 2011.

17. Frederick P. Brooks Jr., *The Mythical Man-Month: Essays on Software Engineering*, anniversary ed.(Boston, MA: Addison-Wesley, 1995; orig. pub.1975), 53．引用的这句话似乎是一句拉丁谚语，被误传为出自古罗马诗人奥维德（Ovid）之口。

18. 这个术语，霍默－狄克逊在《创造力差距》一书中也有使用。

19. Stewart Brand, *The Clock of the Long Now: Time and Responsibility* (New York: Basic Books, 1999), 85.

20. 这种说法源自斯图尔特布兰德的一篇文章。他在文中写道："除了数据格式和数字存储媒介已经消失之外，还存在一个更深层级的问题。大型计算机系统是推动企业、公共机构，以及金融部门发展的核心。随着时间的推移，这些庞大的系统变得异常复杂和不可理解。这是因为新的特性被不断加入，旧的问题通过一层又一层的补丁得到了处理；除此之外，一代又一代的程序员还加入了新的编程工具和风格，系统的某些部分也被重新设计并赋予了新的功能。出于尊敬，也出于厌恶，计算机专业人士称这些'怪物'为'古迹'。从这些被称为'古迹'的系统中引申出新的功能，绝不是一声令下就能完成的事情；只能小心翼翼地进行炼金术般的实验，如果运气好的话，最终也可能会得到想要的结果。"请参阅："Written on the Wind," The Long Now Foundation, February 11, 1998。

21. Christopher Ingraham, "Charted: The Skyrocketing Complexity of the Federal Tax Code," *Washington Post*, April 15, 2015.

22. *Cheek v. United States*, 498 U. S 192.；*United States v. Murdock*, 290 U. S.389.

23. Susan E. Dudley and Jerry Brito, *Regulation: A Primer*, 2nd ed. (Arlington, VA, and Washington, DC: Mercatus Center, George Mason University, and The George Washington University Regulatory Studies Center, 2012), 5.

24. *The Economist*, November 19, 1955．不过，当一个国家处于战争状态时，这种关系似乎并不成立。

25. Meir M. Lehman, “Programs, Life Cycles, and Laws of Software Evolution,” *Proceedings of the IEEE* 68, no.9(1980): 1060-76.Also Lehman et al., “Metrics and Laws of Software Evolution-The Nineties View,” *in METRICS' 97, Proceedings of the Fourth International Software Metrics Symposium* (Washington, DC: IEEE Computer Society, 1997)．另外也请参阅下文中要讨论的约翰·盖尔所著的《系统圣经》。

26. 这种重写其实相当于“重构”，“重构”即重新编写软件代码，使软件的内在更统一、更“干净”，但不一定会改变其外部功能。

27. 宇航员艾伦·谢泼德（Alan Shepard）曾经说过：“虽然身处太空，心里却很清楚，你的安全完全取决于政府招标时出价最低的那个竞标者。一想到这里，不禁心中酸楚。”

28. John Kelly, “Look Out Below: Danger Lurks Underground from Aging Gas Pipes,” *USA Today*, September 23, 2014.

29. Metropolitan Transportation Authority (MTA), “CBTC: Communications-Based Train Control,” July 20, 2015.

30. 而且，为了预测边界情况和例外情况，这些系统甚至不得不在复杂状态下启动，本章下文将对此进行讨论。

31. Harry McCracken, “Fifty Years of BASIC, the Programming Language That Made Computers Personal,” *TIME*, April 29, 2014．对于这类分叉和循环，更严格的控制方法是运用更明确的指令语句，比如for循环语句。

32. Edsger W. Dijkstra, “GoTo Statement Considered Harmful,” *Communications of the ACM* 11, no.3(1968): 147-48．需要注意的是，迪科斯彻这篇论文最初的标题是“A Case Against the GOTO Statement”。

33. 其中一种方法是重构。软件开发者布莱恩·福特（Brian Foote）和约

瑟夫·约德（Joseph Yoder）曾经阐释过，为什么“当前的标准软件构架”其实是“一个松散的，甚至有害的结构体系”，或者说是一个“大泥球”；他们还试图描绘出改进此类系统的内部通道。请参阅：Brian Foote and Joseph Yoder, “Big Ball of Mud,” *Fourth Conference on Pattern Languages of Programs*, Monticello, IL, September 1997; in *Pattern Languages of Program Design 4*, ed. Brian Foote, Neil Harrison, and Hans Rohnert, chapter 29(Boston: Addison-Wesley, 2000)。

34. Koopman, “Case Study of Toyota Unintended Acceleration,” slide 38．库普曼指出，意大利面条式的代码可能是导致高度复杂性的原因之一。总部位于英国的汽车工业软件可靠性联合协会（MISRA）发布了行业标准《MISRA软件指南》，试图提高系统的安全性能。

35. Philip K. Howard, *The Rule of Nobody: Saving America from Dead Laws and Broken Government* (New York: W. W. Norton, 2014).

36. Howard, *The Rule of Nobody*, 8.

37. Howard, *The Rule of Nobody*, 12。以新桥取代旧桥的建设项目通常需要10年时间才能获批立项；至于高速公路建设项目，单是环境评估环节就要花上8年。

38. Michael Mandel and Diana G. Carew, *Regulatory Improvement Commission: A Politically-Viable Approach to U. S. Regulatory Reform*, Progressive Policy Institute policy memo (Washington, DC: Progressive Policy Institute, May 2013).

39. John Palfrey and Urs Gasser, *Interop: The Promise and Perils of Highly Interconnected Systems* (New York: Basic Books, 2012).

40. Sergey V. Buldyrev et al., “Catastrophic Cascade of Failures in Interdependent Networks,” *Nature* 464(2010): 1025–28；Natalie Wolchover, “Treading Softly in a Connected World,” *Quanta Magazine*, March 18, 2013.

41. 失败的成本实际上是一个假想的分布，而不是某个特定的数字；不过，我们也可以将其想象为一个期望值或平均值。

42. J. R. Minkel, “The 2003 Northeast Blackout-Five Years Later,” *Scientific American*, August 13, 2008.

43. 关于构建成本和故障成本的观点，要感谢爱德华·荣格（Edward

Jung）在2014年3月8日的私人信件中给我的启发。

44. “A Maps App with Problems,” *The New York Times*, September 27, 2012.

45. Jeronimo Cello et al., “Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template,”*Science* 297(2002): 1016-18; Eckard Wimmer, “The Test-Tube Synthesis of a Chemical Called Poliovirus: The Simple Synthesis of a Virus Has Far-Reaching Societal Implications,” *EMBO Reports* 7, no.1S(2006): S3–S9.

46. Doug Hill, *Not So Fast: Thinking Twice about Technology* (Cellarius Press, 2013, out of print;Athens, GA: University of Georgia Press, forthcoming);Kevin Kelly, *What Technology Wants*.

47. Dijkstra, “On the Cruelty of Really Teaching Computing Science,” E. W. Dijkstra Archive: The manuscripts of Edsger W. Dijkstra, 1930-2002, document no. EWD1036, December 1988.

48. 系统的复杂性反映了世界的复杂性，请参阅：Vikram Chandra, *Geek Sublime: The Beauty of Code, the Code of Beauty* (Minneapolis: Graywolf Press, 2014)。我们之所以会编写出复杂的代码，并不只是因为这个世界太过复杂，但大多数时候的确如此。幸运的是，有一些方法可以缓解这个问题，请参阅：Steve McConnell, *Code Complete: A Practical Handbook of Software Construction*, 2nd ed.(Redmond, WA: Microsoft Press, 2004), 583。

49. 关于自动驾驶汽车的复杂性，请参阅：Astro Teller, “How to Make Moonshots,” *Backchannel*, March 17, 2015。

50. 一种解决方案是进行人工干预，即手动排除故障，至少要对例外情况进行“硬编码”。谷歌公司正是这样处理谷歌地图的：“这是一种实现超级可靠性的谷歌式（Google-y）方法。谷歌公司旗下许多有名的驱动型计算项目，例如创建地图等，背后都隐藏着数以千计的员工。他们随时监督着系统，并修正系统的错误。这是谷歌公司公开的秘密之一。他们成功地让人工智能变身为催化剂。他们并没有使用编程的方式来构建最后这一道防线，构建这最后的1%或0.1%，甚至0.01%的可靠性；而是通过一部分廉价的人力来实现这一点。未来，人类可以通过指导机器运行，让自己从这项工作中脱身而出。‘如果人类告诉机器，这件事这样做

是正确的，那么这个指令就会成为一个可以融入系统的信息；未来需要人工干预的情况就会逐渐减少。’消息人士这样说。”请参阅：Alexis C. Madrigal, “Inside Google's Secret Drone-Delivery Program,” *The Atlantic*, August 28, 2014。

51. “Friar Daw's Reply,” from *Six Ecclesiastical Satires*, ed. James M. Dean, TEAMS Middle English Texts Series (Kalamazoo, MI: Medieval Institute Publications, 1991)可以从罗切斯特大学罗宾斯图书馆数字项目网站（Robbins Library Digital Projects）上下载。

52. 请注意，包括长尾分布在内，不是所有的重尾分布都是服从幂律的。

53. András Kornai, *Mathematical Linguistics* (London: Springer-Verlag, 2008), 71．研究表明，在多种语料库中，罕见语所占比例约为40%到60%。

54. William A. Kretzschmar Jr., *The Linguistics of Speech* (Cambridge, UK: Cambridge University Press, 2009)．同样相关的，还有弗斯语言学（Firthian linguistics），它“基于这样一种观点：语言模式无法用单一的原则来分析，也无法用某个系统类型来解释”，而必须同时考虑多个与上下文有关联的系统。请参阅：David Crystal, ed., *Dictionary of Linguistics and Phonetics*, 6th ed.(Malden, MA: Wiley-Blackwell, 2008), 181。

55. Peter Norvig, “On Chomsky and the Two Cultures of Statistical Learning”.

56. 这个关于机器翻译的预言般的故事还有好几个版本。

57. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford, UK: Oxf ord University Press, 2014), 15.

58. 这个数字是我自己总结出来的，如果哪位语言学家有不同意见，请告知我。

59. Alon Halevy et al., “The Unreasonable Effectiveness of Data,” *IEEE Intelligent Systems* 24, no.2(2009): 8-12．在某种意义上，和那些乍看起来规则既简单又优雅的模型相比，这些统计学模型其实更加简单，因为前者最后都会因各种始料不及的原因而变得复杂。

60. Douglas Heaven, “Higher State of Mind,” *New Scientist* 219(August

10, 2013), 32-35.

61. Brooks, *Mythical Man-Month*, 183-84．小弗雷德里克·布鲁克斯认识到，复杂性的类型多种多样，甚至包括与软件相互作用的环境所施加的复杂性。

62. David G. Post and Michael B. Eisen, “How Long Is the Coastline of the Law? Thoughts on the Fractal Nature of Legal Systems,” *Journal of Legal Studies* 29, no.2 J(2000): 545-84.

63. Jack M. Balkin, “The Crystalline Structure of Legal Thought,” *Rutgers Law Review* 39, no.1(1986): 1-108.

64. Post and Eisen, “How Long Is the Coastline of the Law?”

65. 波斯特和艾森发现了幂律的基本特性。

66. Mark D. Flood and Oliver Goodenough, “Contract as Automaton: The Computational Representation of Financial Agreements,” OFR(Office of Financial Research) Working Paper no.15-04, March 26, 2015.

67. Duncan Watts, “Too Complex to Exist,” *The Boston Globe*, June 14, 2009.

68. Palfrey and Gasser, *Interop*.

69. 假设每个独立模块能够以3种不同的方式“出站”连接到另一个模块，并以3种不同的方式“入站”连接到其他模块，或者自环连接，那么，我们将会得到大约10^{32}个潜在的网络。在当前人类可观测的宇宙范围中，恒星的准确数量尚不清楚，但据估计，应该不到10^{30}颗。请参阅：Elizabeth Howell, “How Many Stars Are In The Universe?” Space. com, May 31, 2014。

70. 本书的主要论题是如何在一个人类无法理解的时代生存下去，而今，这已经成为一个很现实的问题；尽管如此，我们仍然可以找到一些方法来设计和构建更具可控性的系统，比如“系统思维”这样的方法。请参阅：Nancy G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety* (Cambridge, MA: The MIT Press, 2011)。

71. McConnell, *Code Complete*, 521.

03　为什么复杂系统越来越难以理解了

1. 这个故事和分析取材于：Nancy G. Leveson and Clark S. Turner, “An Investigation of the Therac-25 Accidents,” *Computer* 26, no.7(1993), 18-41。

2. Leveson and Turner, “An Investigation.”

3. 　机器，更准确地说，编程语言也不是不能从“1”开始计数，只不过，如今的大多数编程语言都是从“0”开始计数的。原因非常古老，就连很多程序员都不甚清楚，关于这段历史的记录，请参阅：Michael　Hoye, “Citation Needed,”*blarg? Mike Hoye's weblog*, October 22, 2013。

4. Scott Rosenberg, *Dreaming in Code: Two Dozen Programmers, Three Years, 4, 732 Bugs, and One Quest for Transcendent Software* (New York: Three Rivers Press, 2008), 6-7.

5. Homer-Dixon, *The Ingenuity Gap*, 186.

6. 实际上是这样算的：10 000个名词×1 000个动词×9 999个名词，那么需要3万多年才能说完这些句子。

7. Steven Pinker, *The Language Instinct: How the Mind Creates Language* (New York: William Morrow, 1994; repr. Harper Perennial, 1995), 205.

8. Ray Kurzweil, *The Age of Spiritual Machines: When Computers Exceed Human Intelligence* (New York: Penguin, 1999), 95.

9. Mark Pilgrim, *Dive into Python: Python from Novice to Pro*, updated 2004．虽然“康德发生器”可以将从句多重嵌套进其他句子中，但这个程序似乎不能实现无限递归，因为它总会自动结束。

10. “Garden Path Sentence,” Wikipedia.

11. 　还有一个例子是，我们进行战略性决策时所能考虑到的“步数”。如博弈论所述，当决策依赖于我们所认为的，其他人的想法和行为时，我们就会遇到这个问题。很少有人能够想到很多步，因为大多数人总是根据“他认为我认为他认为我认为……”的原则在做事。请参阅：Colin Camerer et al., “Behavioural Game Theory: Thinking, Learning and Teaching,” Caltech Working Paper。

12. George A. Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information,” *The Psychological Review* 63(1956): 81–97.

13. Lin Zhong, “Limitations of Human Mind,” lecture notes for ELEC513/COMP513, Complexity in Modern Systems, Department of Electrical and Computer Engineering, Rice University.

14. Bostrom, *Superintelligence*, 59-60．对于长期记忆，有一些人给出了很大容量的估值，比如2.5PB。

15. Thomas Hills and Ralph Hertwig, “Why Aren't We Smarter Already: Evolutionary Trade-Offs and Cognitive Enhancements,” *Current Directions in Psychological Science* 20, no.6(2011): 373-77.

16. Jorge Luis Borges, “Funes, His Memory,” in *Collected Fictions*, trans. Andrew Hurley(New York: Viking Penguin, 1998), 131-37.

17. Robert Kanigel, *The Man Who Knew Infinity: A Life of the Genius Ramanujan* (New York: Charles Scribner's Sons, 1991; repr. Washington Square Press, 1992).

18. Chandra, *Geek Sublime*, 48.

19. Jody Rosen, “The Knowledge, London's Legendary Taxi-Driver Test, Puts Up a Fight in the Age of GPS,” *T, The New York Times Style Magazine*, November 10, 2014.

20. 塞萨尔·伊达尔戈（César Hidalgo）创造了个人字节(personbyte)这个术语，用来描述人脑能够容纳的信息量和知识量，请参阅：Hidalgo, *Why Information Grows*。

21. John Symons and Jack Horner, “Software Intensive Science,” *Philosophy and Technology* 27, no.3(2014): 461–77.

22. 莱姆在几十年前就分析过这种不可理解的复杂性，请参阅：Lem, *Summa Technologiae*, trans. Joanna Zylinska(orig. pub. in Polish, 1964; Minneapolis: Unive rsity of Minnesota Press, 2013), 96-97。

23. Jim Gao, “Machine Learning Applications for Data Center Optimization,”Google White Paper.

24. Joe Kava, “Better Data Centers through Machine Learning,” *Google*

Official Blog, May 28, 2014.

25. Douglas Heaven, “Higher State of Mind,” *New Scientist* 219(August 10, 2013), 32-35.

26. 需要注意的是，这种电路实际上是在硬件领域中发展起来的。“种群”中的每个成员都在一个被称为“现场可编程门阵列”（Field-Brogrammable Gate Array）的可编程电路中实例化，而不仅仅只是进行模拟运算。其他许多进化算法则完全发生在软件内部。请参阅：Adrian Thompson, “Exploring Beyond the Scope of Human Design: Automatic Generation of FPGA Configurations Through Artificial Evolution(Extended Abstract),” 8th Annual Advanced PLD and FPGA Conference, 1998。也可参阅：Thompson, “An Evolved Circuit, Intrinsic in Silicon, Entwined with Physics,” in *Evolvable Systems: From Biology to Hardware*, ed. Tetsuya Higuchi et al., Lecture Notes in Computer Science vol.1259(New York: Springer, 2008), 390-405。

27. Kevin Kelly, *Out of Control* (New York: Basic Books, 1994), 338.

28. Steven Rosenbush and Laura Stevens, “At UPS, the Algorithm Is the Driver,” *The Wall Street Journal*, February 16, 2015．在博客文章《边际革命》（*Marginal Revolution*）中，阿历克斯·塔巴洛克（Alex Tabarrok）把这种智能称为“不透明智能”（opaque intelligence）。

29. Tyler Cowen, *Average Is Over: Powering America beyond the Age of the Great Stagnation* (New York: Dutton, 2013), 72.

30. Feng-Hsiung Hsu, *Behind Deep Blue: Building the Computer That Defeated the World Chess Champion* (Princeton, NJ: Princeton University Press, 2002).

31. Flood and Goodenough, “Contract as Automaton.”

32. Donald Sull and Kathleen M. Eisenhardt, *Simple Rules: How to Thrive in a Complex World* (New York: Houghton Mifflin Harcourt, 2015), 12-13.

33. 这三本书分别是：Paula Findlen, ed., *Athanasius Kircher: The Last Man Who Knew Everything* (New York: Routledge, 2004);Andrew Robinson, *The Last Man Who Knew Everything: Thomas Young, the Anonymous Polymath Who Proved Newton Wrong, Explained How We*

See, Cured the Sick and Deciphered the Rosetta Stone (New York: Plume, 2007; orig. pub. Saddle River, NJ: Pi Press/Pearson Education, 2006);Leonard Warren, *Joseph Leidy: The Last Man Who Knew Everything* (New Haven: Yale University Press, 1998)。

34. Philip Ball, *Curiosity: How Science Became Interested in Everything* (Chicago: University of Chicago Press, 2013), 55.

35. Ball, *Curiosity*, 156.

36. Daniel J. Boorstin, *The Discoverers: A History of Man's Search to Know His World and Himself* (New York: Random House, 1983), 414.

37. Boorstin, *The Discoverers*, 414.

38. Ball, *Curiosity*, 120.

39. Ball, *Curiosity*, 120.

40. John M. Ziman, *Knowing Everything About Nothing: Specialization and Change in Research Careers* (Cambridge, UK: Cambridge University Press, 1987),v.

41. Benjamin F. Jones, “The Burden of Knowledge and the ‘Death of the Renaissance Man': Is Innovation Getting Harder?” *The Review of Economic Studies* 76(2009): 283-317．这个理论主要关注的是与技术进步有关的知识。

42. Benjamin F. Jones, E. J. Reedy, and Bruce A. Weinberg, “Age and Scientific Genius,” in *The Wiley Handbook of Genius*, ed. Dean Keith Simonton(Malden, MA: John Wiley and Sons, 2014).

43. “Fascinating Facts,” Library of Congress.

44. Edward O. Wilson, *Consilience: The Unity of Knowledge* (New York: Alfred A. Knopf, 1998; repr. Vintage Books, 1999), 42–43.

45. Jordan Bell-Masterson, “Innovation Series: The Rising Costs of Invention,”*Growthology*, March 24, 2015.

46. Michael Ogawa and Kwan-Liu Ma, “Software Evolution Storylines,” *SOFTVIS' 10: Proceedings of the 5th International Symposium on Software Visualization* (New York: ACM Digital Library, 2010), 35-42.

47. Brooks, *Mythical Man-Month*, 180.

04　令人费解的bug

1.　“但是，这几十年来，在我的脑海中始终有一些问题挥之不去。这个作弊法是故意被添加到代码中，让内线玩家作为后门来使用的吗？或者只是软件中的一个小故障，是连续几个版本的游戏只读存储卡中一直存在的某个意外问题所引发的副作用？如果这个作弊法只是一个小故障，那么代码到底出了什么问题？是否有其他动作序列可以更快地到达禁用状态？”请参阅：Chris Cantrell, “Arcade / Galaga,” Computer Archeology。我还了解到，不少人认为这是一个有意为之的骗局，例如下面这篇文章，就持有这种观点：Jason Eckert, “The Galaga No Fire Cheat Mystery,” October 31, 2012(updated May 2014)。我第一次知道《小蜜蜂》有这个问题，有赖于这篇文章：Clive Thompson, *Smarter Than You Think: How Technology Is Changing Our Minds for the Better* (New York: Penguin, 2013)。

2. 谷歌公司杰出的研究员乌尔斯·霍尔泽(Urs Hölzle)指出：“对于致力于推进宏伟计划的人来说，复杂性在某种意义上是邪恶的，因为它会让某些bug两三年才冒出来一次；而当你看到这些bug时，就意味着已经出了大事，因为它具有巨大的级联效应。”请参阅：Jack Clark, “Google: ‘At Scale, Everything Breaks,’” *ZDNet*, June 22, 2011。

3. A. M. Turing, “Computing Machinery and Intelligence,” *Mind* 59(1950): 433-60.

4.　在3万行代码中，每千行就会产生大约1万个特定类型的错误。请参阅：John Symons and Jack Horner, “Software Intensive Science,” *Philosophy and Technology* 27, no.3(2014): 461-77; J. K. Horner, “Persistence of Plummer-Distributed Small Globular Clusters as a Function of Primordial-Binary Population Size,” *Proceedings of the International Conference on Scientific Computing(CSC)* (Athens: CSREA Press, 2013), 100-106。

5. 有人估计，每3至5行代码就存在一个错误，这个数字无疑是令人吃惊的。请参阅：Roger A. Grimes, “In His Own Words: Confessions of a Cyber Warrior,” *InfoWorld*, July 9, 2013。

6. Bruce Brown, Nigel R. Smith, and Bruce Kratofil, *The Windows 95 Bug Collection* (Reading, MA: Addison-Wesley, 1996), 3-10.

7. McConnell, *Code Complete*, 652.

8. J. M. Carlson and John Doyle, “Complexity and Robustness.” *PNAS* 99, Suppl.1(2002): 2538-45.

9. Carlson and Doyle, “Complexity and Robustness.”

10. Koopman, “Case Study of Toyota Unintended Acceleration,” slide 20.

11. 这些著作包括：*The Words*： *Anatomy of a City* (2005)；*The Heights: Anatomy of a Skyscraper* (2011); *The Way to Go: Moving by Sea, Land, and Air* (2015)。这几本著作均由企鹅出版集团出版。

12. “Massachusetts Water Crisis,” *The Boston Globe*.

13. Andrew Blum, *Tubes: A Journey to the Center of the Internet* (New York: Ecco, 2012).

14. Roger A. Grimes, “Shellshock Proves Open Source's ‘Many Eyes’ Can't See Straight,” *InfoWorld*, September 30, 2014.

15. Jon Brodkin, “Why Gmail Went Down: Google Misconfigured Load Balancing Servers(Updated),” *Ars Technica*, December 11, 2012.

16. Joshua Bloch, “Extra, Extra-Read All About It: Nearly All Binary Searches and Mergesorts Are Broken,” *Google Research Blog*, June 2, 2006．在以下论著中也有相关讨论：Chandra, *Geek Sublime*, 124。

17. 一个小故障是“浏览软件内部结构的可能性”，请参阅：Olga Goriunova and Alexei Shulgin, “Glitch,” in *Software Studies: A Lexicon*, ed. Matthew Fuller, 110-19 (Cambridge, MA: The MIT Press, 2008), 114。错误和bug可以为改进软件提供机会，可参阅纳西姆·尼古拉斯·塔勒布（Nassim Nicholas Taleb）的论著：*Antifragile: Things That Gain from Disorder* (New York: Random House, 2012)。关于塔勒布的“反脆弱”思想在软件开发中的应用，请参阅：Martin Monperrus, “Principles of Antifragile Software”。

18. 在很多地方都可以看到与温哥华证券交易所指数事件有关的故事。请参阅：Anne Greenbaum and Timothy P. Chartier, *Numerical*

Methods: Design, Analysis, and Computer Implementation of Algorithms (Princeton, NJ: Princeton University Press, 2012), 117; Kevin Quinn, "Ever Had Problems Rounding Off Figures? This Stock Exchange Has," *Wall Street Journal*, November 8, 1983。

19. "混沌猴"可以在网上下载。

20. Bill Bryson, *A Short History of Nearly Everything* (New York: Broadway Books, 2003), 162.

21. George Dyson, *Turing's Cathedral: The Origins of the Digital Universe* (New York: Pantheon, 2012), 4.

05 为什么需要生物学思维

1. *Dictionary of National Biography, 1885–1900*, vol.18, "Fairfax, Nathaniel, M. D."需要注意的是，这位医生发表论文时的署名，在拼写上不尽相同，但是通过其传记资料，可以证明是同一个人。

2. Nathanael Fairfax, "Divers Instances of Peculiarities of Nature, Both in Men and Brutes;Communicated by the Same," *Philosophical Transactions* 1666-67, 2(1666): 549-51.

3. 在瘟疫肆虐的那些年，牛顿有时也会在剑桥大学从事各种科学研究。关于牛顿的生平和著作，可以参考的书很多，例如：George Smith, "Isaac Newton,"*The Stanford Encyclopedia of Philosophy*, Fall 2008 edition, ed. Edward N. Zalta; V. Frederick Rickey, "Isaac Newton: Man, Myth, and Mathematics," *The College Mathematics Journal* 18, no.5(1987): 362-89。

4. Tania Lombrozo, "Must Science Murder Its Darlings?" *NPR 13.7: Cosmos and Culture*, January 27, 2014.

5. Freeman J. Dyson, *Infinite in All Directions*, repr. ed.(New York: Harper Perennial, 2004; orig. pub.1988), 40.

6. Shane Parrish, "What Made Charles Darwin an Effective Thinker? Follow the Golden Rule," *Farnam Street*, January 11, 2016．除了科学领域，还有其他一些领域也拥有这种生物学倾向。例如，历史学家有时会有意远离抽象化和一般概念，以便让自己能更好地理解历史的复杂

性。

7. 如果说生物学家和物理学家一样，也在拓展“实验还原主义”的极限，那么他们也会变得和工程师一样，沉迷于所研究系统的庞大规模和多样性，以及纯粹的复杂性。他们对“球形奶牛”这种概念完全没有兴趣。请参阅：John Doyle, “Computational Biology: Beyond the Spherical Cow,” *Nature* 411(2001): 151-52。

8. Steven A. Benner, “Aesthetics in Synthesis and Synthetic Biology,” *Current Opinion in Chemical Biology* 16, no.5-6(2012): 581-85.

9. 过时的遗留代码类似于软件中的“讨厌的遗迹”，对于当前版本的软件来说，这些东西其实是不必要的，但它们存在的时间却会比我们所希望的要久。

10. Whit Bronaugh, “The Trees That Miss the Mammoths,” *American Forests*, Winter 2010.

11. 过时的遗留代码不一定会被自然淘汰，但是可以通过对代码的定期整理和清扫来移除，这一点与生物学规律类似。

12. 乔治·戴森使用了另一个术语：数字生物学家，请参阅：“A Universe of Self-Replicating Code,” *Edge*.

13. “RNAi,” *NOVA*, July 26, 2005.

14. Howard Wainer and Shaun Lysen, “That's Funny,” *American Scientist* 97, no.4(2009): 272.

15. Dan Hurley, “Why Are So Few Blockbuster Drugs Invented Today?” *The New York Times Magazine*, November 13, 2014．我们可以从药品的临床试验中学到很多东西，这一点是爱德华·荣格提醒我的。

16. Stewart Brand, *The Clock of the Long Now: Time and Responsibility* (New York: Basic Books, 1999), 85.

17. Peter G. Neumann, *Computer-Related Risks* (New York: ACM Press, 1995), 122.

18. Neal Stephenson, *Cryptonomicon* (New York: Avon Books, 1999; repr.2002), 802-3.

19. 需要注意的是，在《范布伦家的男孩们》（*The Van Buren Boys*）那一集中，酒吧里还有一个人也被称为“拉米雷斯”，不过我认为那个名字

的发音明显不同于科尔基·拉米雷斯的“拉米雷斯”。也许，他当时就在那，但这只是我的猜测。《宋飞正传》的粉丝，请发邮件告诉我内情！

20. 也可以写为“Greebles”，请参阅：Kelly, *What Technology Wants*, 318。

21. Benoit B. Mandelbrot, *The Fractal Geometry of Nature* (New York: W. H. Freeman and Company, 1982), 1.

22. Borges, “Funes, His Memory,” in *Collected Fictions*, 131–37.

23. Philip Ball, *Branches*, vol.3 of *Nature's Patterns: A Tapestry in Three Parts* (Oxford, UK: Oxford University Press, 2009), 181.

24. William Li et al., “Law Is Code: A Software Engineering Approach to Analyzing the United States Code,” *Journal of Business and Technology Law* 10, no.2(2015): 297-372.

25. Jonathan Barnes, ed. and trans., *Early Greek Philosophy* (London: Penguin Classics, 1987); *The Stanford Encyclopedia of Philosophy*.

26. Barnes, *Early Greek Philosophy*, 72.

27. 这里关于“宇宙”和“本原”的观点都源于：Barnes, *Early Greek Philosophy*。

28. Ball, Curiosity, 98-99.

29. Lorraine Daston, “The Language of Strange Facts in Early Modern Science,”*Inscribing Science: Scientific Texts and the Materiality of Communication*, ed. Timothy Lenoir(Stanford, CA: Stanford University Press, 1998), 20-38.

30. Daston, “Language of Strange Facts,” 38.

31. Johan Bollen et al., “Clickstream Data Yields High-Resolution Maps of Science,” *PLoS ONE* 4, no.3(2009): e4803.

32. Nicholas Donofrio, Jim Spohrer, and Hossein S. Zadeh, “Research-Driven Medical Education and Practice: A Case for T-shaped Professionals,” *MJA Viewpoint*, 2009.

33. 对于专业化的深化发展趋势，霍默 - 狄克逊曾表示对“专家的狭隘化”十分担忧，请参阅：Homer-Dixon, *The Ingenuity Gap*, 176。

34. David J. Teece, “A Dynamic Capabilities Perspective on Building Firm-Level Competitiveness,” slide 43, Tusher Center on Intellectual Capital.

35. Nicola Griffith, *Hild: A Novel* (New York: Farrar, Straus and Giroux, 2013).

36. John Ptak, “A Cloud Map(1873),”JF Ptak Science Books, January 13, 2015.

06　生物学思维是理解复杂世界的一把金钥匙

1. Moses Maimonides, *The Guide of the Perplexed*, vol.1, trans. Shlomo Pines(Chicago: The University of Chicago Press, 1963), 65-66.

2. J. B. S. Haldane, *Possible Worlds and Other Essays* (London: Chatto &Windus, 1928), 286.

3. 关于“科学谦卑”的深入讨论，请参阅：Marcelo Gleiser, *The Island of Knowledge: The Limits of Science and the Search for Meaning* (New York: Basic Books, 2014)。

4. Ian Bogost, “The Cathedral of Computation,” *The Atlantic*, January 15, 2015.

5. Bogost, “Cathedral of Computation.”

6. Edsger Dijkstra, “The Humble Programmer.” *Communications of the ACM* 15, no.10(1972): 859-66.

7. David Brooks, *The Road to Character* (New York: Random House, 2015), 263.

8. Carl Zimmer, “Is Most of Our DNA Garbage?” *The New York Times Magazine*, March 5, 2015.

9. 正文中所引用的这些例句都来自盖尔所著作品的附录I：John Gall, *The Systems Bible: The Beginner's Guide to Systems Large and Small*, 3rd ed.(Walker, MN: The General Systemantics Press, 2003)。

10. 盖尔甚至建议他的读者“珍藏你的系统故障”，就像生物学家收集有趣的观察结果一样；他还以达尔文为例，强调我们必须承认“大型复杂系

统是超出人类评估能力的”。请参阅：*System Bible*, xx, 68。

11. Robert Herritt, “When Technology Ceases to Amaze,” *The New Atlantis* 41, Winter 2014, 121-31.

12. Donald A. Norman, *Living with Complexity* (Cambridge, MA: The MIT Press, 2010), 117-18.

13. Daniel Engber, “Who Made That Progress Bar?” *The New York Times Magazine*, March 7, 2014.

14. Kate Greene, “How Should We Program Computers to Deceive?” *Pacific Standard*, September 3, 2014.

15. 菲利普·郭（Philip Guo）曾写道：“我想知道特波税务软件的代码中到底嵌套了多少if语句，”感谢丹·卡茨的提醒，特波税务软件可以被视为税法的一个界面。

16. Chaim Gingold, *Miniature Gardens and Magic Crayons: Games, Spaces, and Worlds*, master's thesis, Georgia Institute of Technology, 2003, 62.

17. Nicholas Carr, *The Glass Cage: How Our Computers Are Changing Us* (New York: W. W. Norton, 2014).

18. Winner, *Autonomous Technology*, 285.

19. Eytan Adar et al., “Benevolent Deception in Human Computer Interaction,”*CHI’13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,*Paris, France, April 27—May 2, 2013(New York: ACM Digital Library, 2013): 1863-72.

20. 关于《模拟城市》的细节以及它对我们的启发，请参阅：Doug Bierend, “SimCity That I Used to Know: On the Game's 25th Birthday, a Devotee Talks with Creator Will Wright,”re: form, October 17, 2014。

21. Cowen, *Average Is Over*, 227–28．考恩的预言所针对的是经济学以及其他社会科学的未来，但在复杂性科学领域，同样需要能够解释未来社会的人。

22. 正如考恩所说，这些解释者在未来很可能会与机器通力合作。我们可能会喜欢上这种“人首马身”的怪物，当然，那不是半只马，而是半个

机器。请参阅：Clive Thompson, *Smarter Than You Think: How Technology Is Changing Our Minds for the Better* (New York: Penguin, 2013)。

23. David Cope, “Experiments in Musical Intelligence.” Ryan Blitstein, “Triumph of the Cyborg Composer,” *Pacific Standard*, February 22, 2010.

24. 戴维·科普甚至在他的个人网站上写下了如下一段话，将这种“父与子”的自豪感扩展到了全人类的范畴：“算法所创作的音乐，与最具灵感的音乐家所创作的音乐一样，都是人类杰出的作品。”请参阅：Cope, “Experiments”。而且，他还说过这样的话：“计算机就是我自己的延伸。它们不是别的，就是一些非常好的铲子。我不会把挖好洞的功劳归于铲子，你会吗？”请参阅：Blitstein, “Triumph of the Cyborg Composer”。

25. 机器人专家汉斯·莫拉维克（Hans Moravec）把“更强大的后代”称为“心智儿童”(mind children)，在科幻作家姜峯楠（Ted Chiang）的一篇短篇小说中也有类似的情节设定。在这篇小说中，通过科技手段被强化的人，在科学探索方面的能力远远超过了普通人。最终，未被强化的，也就是普通人，则沦落到一无所知的地步。但是没关系，因为“我们用不着对‘超人’的科学成就感到恐惧。我们应该时刻牢记，构建出‘超人’的技术，最初是由人类开发的，所以它们并不比我们聪明”。请参阅：Luke Muehlhauser and Nick Bostrom, “Why We Need Friendly AI,” *Think* 36, no.13(Spring 2014), 41-47; Ted Chiang, *Stories of Your Life and Others* (New York: Tor Books, 2003), 203。

26. *The World of Wonders: A Record of Things Wonderful in Nature, Science, and Art* (London: Cassell, Petter, and Galpin, exact year of publication unknown).

27. 关于如何应对那些避重就轻的故事，请参阅：Philip Ball, “The Story Trap,” *Aeon*, November 12, 2015。

28. MLB. com, “Official Rules: 2.00 Definition of Terms.”

29. 这只是我个人对这两个术语的理解，毫无疑问，还有其他的区分方法。

30. Hubert Dreyfus and Sean Dorrance Kelly, *All Things Shining: Reading the Western Classics to Find Meaning in a Secular Age* (New

York: Free Press, 2011), 88.

31. 关于如何克服对技术的无知，请参阅：Herritt, “When Technology Ceases to Amaze”。

32. “Customer FAQs Regarding the Sticking Accelerator Pedal and Floor Mat Pedal Entrapment Recalls,” Toyota Pressroom.

33. 例如，我们现在还无法完全理解自动驾驶汽车的所有技术细节，但这并不意味着，这种汽车比人类驾驶的汽车更危险。而且，对于人类驾驶的汽车，我们的理解也并不透彻，更不用说开车的人了。

34. 我要感谢提出这个建议的人，还要感谢堪萨斯大学哲学系的同事们，是他们为我提供了限制性定理的相关思路。

35. Carlin Romano, *America the Philosophical* (New York: Alfred A. Knopf, 2012), 501.

延伸阅读书目

本书正文未能呈现本人思想的全部来源。我在书中所阐述的思想，无论是来源还是基础，都很广泛。除了书中所提到的那些论著之外，我还想在这里强调一些引人深思的著作、论文、书籍和文章。在我撰写本书时，它们对我帮助不小。虽然有些著作和论文在正文中未被提及，但是如果你有兴趣深入研究本书所讨论的各个主题，那么这些参考文献应该都会很有用。

此外，限于篇幅，还有很多相关的重要主题没能在本书中被详细讨论到。它们包括："技术变革的发生发展及其对教育的影响"，"大数据"，"人类与机器日益亲密的'合作伙伴'关系"，"自动化趋势和未来就业之间的关系"，等等。另外，我们身边的各种不同类型的系统也很值得分析和讨论，例如制造业、食品业、政府部门、能源行业等。下面这些论著分析了上述主题与系统。

约翰·盖尔的《系统圣经》收录了一系列关于大型系统的奇闻轶事，皆与系统运行与系统故障有关。这是一本很有趣的书，我的许多想法都与盖尔的观点不谋而合。

凯文·凯利的《失控》阐释了很多生物学思维，并讨论了技术如何变得越来越生物化且无法被完全理解。该书观点与本书类似，不过是从涌现和生物复杂性，以及"如何使用生物学原理来构建技术"的角度来进行阐释的。此外，我还要推荐凯利的另一本著作《科技想要什么》（*What Technology Wants*）。

兰登·温纳的《自主的技术》，该书的倒数第2章特别值得一看，里面涉及的一些问题也正是本书所讨论的问题。

托马斯·霍默 - 狄克逊的《创造力差距》讨论了我们应该如何应对日趋复杂和高深莫测的世界。该书还描述了各种各样的系统，其中有一些是本书未涉及的，它们都相当令人着迷。

查尔斯·佩罗（Charles Perrow）的《正常事故》（*Normal Accidents*）是一本关于系统失败和"如何与高风险技术共存"的经典论著。

维克拉姆·钱德拉的《极客升华》（*Geek Sublime*）对编程和计算机

的本质进行了绝妙的反思，涉及了很多迷人的主题。

弗里曼·戴森的《全方位的无限》（*Infinite in All Directions*）很好地阐释了科学思想和发现的本质。

小弗雷德里克·布鲁克斯的《人月神话》是探讨软件开发与设计的经典作品。它重点讨论了工程师团队的管理问题，同时在软件和编程的性质方面提出了很多有趣的想法。

菲利普·鲍尔的《好奇心》（*Curiosity: How Science Became Interested in Everything*）非常出色地解释了科学家，特别是早期的科学家是如何探索世界，如何在极端多样化、极端繁杂的世界中寻找秩序感的。

唐纳德·诺曼的《与复杂性共存》（*Living with Complexity*）从设计的角度出发，考察了复杂性的起源，以及复杂化趋势的必然性。

布雷登·R.艾伦比（Braden R. Allenby）和丹尼尔·萨雷维茨（Daniel Sarewitz）共同撰写的《技术 - 人类条件》（*Techno-Human Condition*）探讨了我们应该如何应对行将到来的技术变革。它对“邪恶的复杂性”的分析特别有意思。

尼克·波斯特洛姆（Nick Bostrom）的《超级智能》（*Superintelligence*）探讨了与超智能机器发展有关的诸多问题。

凯特·阿舍尔的《建筑》（*The Works*）、《高度》（*The Heights*）与《未来的方向》（*The Way to Go*）主要考察了城市、摩天大楼和交通网络，并分析了它们在现实世界中的运转方式，都是极具感染力的令人着迷的书。

埃里克·布林霍夫森（Erik Brynjolfsson）和安德鲁·麦卡菲（Andrew McAfee）合著的《第二个机器时代》（*The Second Machine Age*）考察了我们正在经历的、被寄予厚望的、高速发展的技术变革时代，并探讨了变革对经济的影响，以及我们应该如何应对等主题。

尼古拉斯·卡尔（Nicholas Carr）的《玻璃笼子》（*Glass Cage*）一书中阐述了自动化的危险性，以及我们身边的复杂技术。

马修·克劳福德（Matthew B. Crawford）的《摩托车修理店的未来工作哲学》（*Shop Class as Soulcraft*）[16]中探讨了重新“亲近”技术的重要性。亲力亲为，是美德的一部分。

由斯坦尼斯瓦夫·莱姆（Stanisław Lem）撰写，乔安娜·齐琳斯卡

（Joanna Zylinska）翻译的《科技全书》（*Summa Technologiae*）讨论了自20世纪60年代以来，科技领域内出现的各种问题。这本出自未来科幻小说家之手的著作，特别强调了人类理解能力的极限。

迈克尔·莫布森（Michael Mauboussin）的《三思而后行》（*Think Twice*）讨论了如何才能对身边的复杂系统做出正确的思考。莫布森指出，正确的思考往往是违反直觉的。

罗伯特·赫里克（Robert Herritt）在《新亚特兰蒂斯》（*New Atlantis*）2014年冬季号，即总第41期，121—131页发表的《当技术不再令人惊奇》（*When Technology Ceases to Amaze*），是一篇讨论技术奇迹、复杂性和惊异感的优秀论文。

尼尔·斯蒂芬森的论文《最初……一切只是命令行》（*In the Beginning...Was the Command Line*），后来扩展为一本同名小册子，详细地探讨了计算的协调问题。尽管书中的观点在现在看来已经有点过时了，但在关于“人类如何与技术相联系，又如何与技术分离”这个方面，还是颇具见地的。

伦纳德·E.里德（Leonard E. Read）的《铅笔的故事》（*I, Pencil*）是一篇隽永的短文，它探讨了制造一支铅笔所涉及的高度互联的社会经济体系。对于这个系统，任何人都无法完全理解。

我还建议读者阅读一下艾兹格·迪科斯彻的其他论著。他的许多作品都是计算机科学领域的经典，同时也涉及了有关技术本质的哲学思考。

如需更多阅读材料，请访问我的个人网站，查找相关主题的论文。

致谢

感谢我的编辑妮基·帕帕佐普洛斯（Niki Papadopoulos），她对本书的写作给予了宝贵的支持。感谢马克斯·布罗克曼（Max Brockman），他帮我提炼了本书的核心思想。

本书的部分内容曾在其他地方发表过，包括《万古》（*Aeon*）、《鹦鹉螺》（*Nautilus*）、《页岩》、《连线》等杂志，以及Edge网站出品的《如何思考会思考的机器》（*What Do You Think About Machines That Think?*）一书。我为能有机会公开阐述个人思想而表示感谢。我还要感谢《连线》杂志的科学栏目邀请我开通了博客，我在本书中提到的不少想法都曾以更为不连贯的形式在博客上发表。此外，感谢我电子邮件简报的订阅者们，他们也提前读到了本书的若干内容。

感谢我所有的读者和对谈者，感谢大家慷慨付出的大量时间和丰富的专业知识，在今天这个时代，这两者都非常珍贵。他们是：乔希·阿贝斯曼（Josh Arbesman）、泽夫·伯杰（Zev Berger）、安德鲁·布卢姆、亚伦·克劳塞特（Aaron Clauset）、洛丽·埃默生（Lori Emerson）、乔舒亚·费尔菲尔德（Joshua Fairfield）、亨利·法雷尔（Henry Farrell）、劳伦斯·冈萨雷斯（Laurence Gonzales）、爱德华·荣格、亚伦·卡恩（Aaron Kahn）、丹·卡茨（Dan Katz）、迈克尔·科兴德费尔（Mykel Kochenderfer）、史蒂文·米勒（Steven Miller）、梅根·欧文（Megan Owen）、埃尔纳坦·赖斯纳（Elnatan Reisner）、内厄姆·沙尔曼（Nahum Shalman）、雅各布·舍曼（Jacob Sherman）、特德·斯坦伯格（Ted Steinberg）、布赖恩·斯蒂芬斯（Brian Stephens）、哈里·萨尔顿（Harry Surden）和杰文·韦斯特（Jevin West）。上面这些人，以及其他一些人，帮助我对本书进行了修订，并避免了一些相当荒谬的错误；也就是说，书中若还存在错误，责任都在我自己身上。

我还要特别感谢堪萨斯大学哲学研究小组的所有成员。该小组致力于考察人类的理解能力与软件之间的关系。作为堪萨斯大学哲学系的一名访问学者，我也是这个小组的成员之一。小组成员约翰·西蒙斯、杰克·霍纳和拉蒙·阿尔瓦拉多（Ramon Alvarado）的意见和建议，对本书而言具有无法估量的价值。戴维·斯蒂恩（David Steen）为我提供了有关野生动物生态学和野外生物学方面的指引；乔希·森夏恩（Josh Sunshine）在计算机科学和软件工程方面给了我大量的支持和建议；迈

克尔·维特维奇（Michael Vitevitch）帮助我厘清了语言的本质和人类处理语言的方式等方面的问题；迈克尔·巴尔和菲利普·库普曼就丰田汽车意外加速事故和技术的复杂性问题，为我提供了有益的建议。

感谢尤因·马里恩·考夫曼基金会（Ewing Marion Kauffman Foundation），它为我创造了一个很好的写作环境。感谢基金会的同事们，感谢你们无与伦比的热情和快速的反馈，特别感谢你们的大度，允许我按自己的进度完成本书。

此外，拉克斯资本（Lux Capital）的项目团队也对本书的写作给予了极大的支持，并为我提供了大量的建议和意见。非常感谢大家。

最后，感谢我的家人。我的父母在第一时间为我提供了支持和建议，并针对初稿给出了详细意见。祖父是我最出色的听众。我非常感谢家里的每一个人。特别感谢我的妻子德布拉（Debra），是你给了我空间和时间，让我可以保质保量地完成这本书。在整个写作过程中，从创作初稿开始，你的不急不躁一直令我惊叹不已。

感谢阿比高尔。你是那么喜欢读书，尽管现在的你更喜欢看那些文字简单、色彩炫丽的书。你每天都会为我带来无限欢乐。从第一行字开始，你一直支持我写这本书，尽管你的支持主要表现为热情的拥抱、无拘无束的欢笑和快乐的闲聊，但是这些已经比我所期望的要多得多了。虽然我也希望能有众多的读者喜欢这本书，但它终究是为你和你的弟弟内森写的。你们俩注定要在一个日益复杂的世界中长大成人。未来世界会是什么样，现在的我只能靠想象。不过，我敢肯定的是，它与我们今天所处的世界大为不同，它将变得更加难以理解。我希望通过这本书，以一种举重若轻的方式让你明白，为什么你们不用为那样的未来感到害怕。

译者后记

在这个越来越复杂的世界上，我们已经无法理解自己创造出来的系统了。技术系统、经济系统、政治系统、法律系统，莫不如是。这是一个悖论。这些过于复杂的系统，在使我们的生活更加便利的同时，本身却变得无法把控了；更加严重的问题是，我们无法预知，这些系统会在什么时候突然崩溃，造成不可承受的损失。

从技术的角度来看，这种复杂性产生于为了适应边界情况而不断加入补丁的过程。任何一个人造系统，哪怕最初是根据清晰的数学模型构建，用于非常明确的用途，只要一直被使用，就肯定会成为一个补丁加补丁的拼凑起来的系统。这是一个进化过程，让人造系统越来越远离确定性。

那么应该怎么办呢？塞缪尔·阿贝斯曼告诉我们，虽然我们应该努力去理解人类无法理解的复杂系统，但首先必须接受混乱。只有这样，我们才能静下心去观察各种意想不到的事件，获得关于我们所用的算法是怎样真正起作用的线索。他强调，要认识世界的复杂性，必须要结合物理学和生物学：用生物学思维处理个例，用物理学思维提取规律。总之，要承认复杂性的存在，同时不要放弃理解系统的原理，而且要从简单的组件开始。

塞缪尔·阿贝斯曼在这本书中给我们提供了应对超复杂这个难题的一些方法，但是并没有提供终极答案。事实上，这个问题也不应该有终极答案。这是因为，复杂是人类行动的非意图后果的集中反映。一方面，在进入现代社会后，我们的任何一个行动，都可能产生比以往任何时候都意想不到的后果；另一方面，我们对问题的暂时解决，反而可能掩盖了问题的根本原因。但是，这种“困境”可以说是不可避免的。它虽然有时会令我们承受痛苦，但却可能是人类进步最主要的途径。它也告诉我们，一劳永逸解决问题的可能性从来就是微乎其微的。这一点对公共政策领域的决策者尤其重要。

说到底，思考和应对复杂问题，需要的是思维，以及思维的多样性。一言以蔽之，真正重要的是，我们要努力让自己成为创新者、思想家和科学家。

这本书的翻译得以完成，要感谢很多人的支持和帮助。

我首先要感谢的是我的太太傅瑞蓉。我的每一本译著，她都有一大半的功劳。同时也要感谢儿子贾岚晴带给我的动力和快乐。

感谢有情怀、重情义、善创新、爱读书的农夫山泉董事长钟睒睒，他既是好老板，也是好老师，让我在工作之余能够腾出时间来完成此书的翻译。

还要借此机会感谢汪丁丁教授、叶航教授和罗卫东教授的教诲。我对复杂性科学的兴趣源于他们。同时还要感谢何永勤、虞伟华、余仲望、鲍玮玮、傅晓燕、傅锐飞、陈叶烽、童乙伦、罗俊、邓昊力、陈姝、黄达强、李燕、李欢、丁玫、何志星等好友的支持和帮助。

感谢湛庐文化。感谢简学。

书中错漏之处在所难免，敬请专家和读者批评指正。

贾拥民

于杭州耑谷阁

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读APP：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”APP。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读APP玩转指南

湛庐阅读APP结构图：



三步玩转湛庐阅读APP：



使用APP扫一扫功能，
遇见书里书外更大的世界！



快速了解本书内容，
湛庐千册图书一键购买！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

延伸阅读

《技术的本质》经典版

- ◎ 复杂性科学的奠基者、著名的技术思想家、“熊彼特“奖的得主布莱恩·阿瑟的经典作品。
- ◎ 埃里克·施密特、凯文·凯利、汪丁丁、段永朝、陈劲、包国光等联袂推荐。
- ◎ 技术理论体系的先河之作，独具创新的关于技术产生和进化的系统性理论！一次打开“技术黑箱”的尝试性创新探索！



使用“湛庐阅读”APP，
“扫一扫”获取本书更多精彩内容。



《复杂经济学》

- ◎ 作为“复杂经济学”的创始人，布莱恩·阿瑟在本书中汇集了多年对复杂经济学的研究。其核心思想可以归结为：经济不一定处于均衡状态，演绎推理将被归纳推理所取代。
- ◎ 这是一本见证“复杂经济学”成长的著作，你将看到“复杂性思维”在经济学领域是如何发展起来并形成一门崭新的学科。
- ◎ 详细讨论了“爱尔法鲁酒吧问题”和“圣塔菲人工股票市场”两个重要项目。





《多样性红利》

◎ 一个人是否聪明不是由智商决定的，而取决于认知工具的多样性！

◎ 《多样性红利》创造性提出多样性视角、启发式、解释和预测模型四个认知工具箱框架。

◎ 广受欢迎的“模型思维课”主讲人斯科特·佩奇经典名作！





《同步：秩序如何从混沌中涌现》

◎ “复杂性理论”开创性研究者斯蒂芬·斯托加茨重磅新作！

◎ 互联网时代最重要的观念之一“小世界网络“的提出者！

◎ 电子科技大学教授周涛，诺贝尔奖获得者菲利普·安德森，牛津大学教授罗伯特·梅，美国海军研究实验室卢·佩科拉等联袂推荐！



使用“湛庐阅读”APP，
“扫一扫”获取本书更多精彩内容。



(1) 即隐藏的代码，通常用于实现某些系统功能或提示某些界面信息等。——编者注

(2) 意为被设计得过度复杂的机械组合，以迂回曲折的方法完成一些其实非常简单的工作。——编者注

(3) 指将数据直接嵌入到程序或其他可执行对象的源代码中的软件开发实践。——编者注

(4) 幂律来自20世纪20年代对英语单词使用频率所做的分析，语言学家发现单词使用的频率和它的使用优先度是一个常数次幂的反比关系。它拥有两个通俗化的解释，一个是长尾理论，另一个是马太效应。——编者注

(5) 又称拉普拉斯-高斯曲线，或正态分布曲线，两端低中间高，常被数学家用来描述科学观察中量度与误差的分布。——编者注

(6) 史蒂芬·霍金在《时间简史》里讲过这样一个故事：一位科学家在演讲结束时被一位老妇人讥讽：“你说的都是废话。这个世界不过是一个乌龟背上的平板。”科学家问：“那么，这只乌龟是站在什么上的呢？”老妇人说：“当然是乌龟群啊！”——编者注

(7) 这些应用程序的功能大体相同，但是运行方式不同，例如微软系统中的PowerPoint与苹果系统中的Keynote。

(8) 史蒂芬·平克的著作《当下的启蒙》《思想本质》《白板》《语言本能》《心智探奇》已由湛庐文化策划，浙江人民出版社出版。——编者注

(9) 爱德华·威尔逊的著作《创造的本源》《半个地球》等，已由湛庐文化策划，浙江人民出版社出版。——编者注

(10) 研究“生物各层次结构与功能的关系”“生命活动的物理”“物理化学过程”和“物质在生命活动过程中表现的物理特性”的生物学分支学科。——编者注

(11) 在管理工作中运用线性规划、非线性规划、动态规划和整数规划，以及系统科学方法所确定的表示最优方案的模型。——编者注

(12) 乔治·戴森的著作《图灵的大教堂》已由湛庐文化策划，浙江人民出版社出版。——编者注

(13) 是一种概率分布模型，包含长尾分布和次指数分布两种子类型。——编者注

(14) 即失去原本矮壮的造型，茎叶疯狂伸长的现象。主因是缺少日照，或光线过暗，同时浇水又相对较多。——编者注

(15) 美国环保主义动画影集《地球超人》中的人物。——编者注

(16) 本书已由湛庐文化策划，浙江人民出版社出版。——编者注

与思想有关
business

THE LIFE SCIENCES 生命科学书系

Arrival of the Fittest

适者降临

自然如何创新

[美]安德烈亚斯·瓦格纳◎著
(Andreas Wagner)
祝锦杰◎译



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

Arrival of the Fittest

适者降临

［美］安德烈亚斯·瓦格纳◎著

（Andreas Wagner）

祝锦杰◎译

浙江人民出版社

ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2018年2月出版

由Current, a member of Penguin Group（USA）LLC授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：适者降临

作者：安德烈亚斯·瓦格纳

译者：祝锦杰

字数：207千

电子书定价：65.99元

Arrival of the Fittest : Solving Evolution's Greatest Puzzle.

Copyright © 2014 by Andreas Wagner.

进化论的最大谜题：自然如何创新？

1．“基因”（gene）这个概念的提出者是下列哪位？（　　）

A．达尔文

B．孟德尔

C．德弗里斯

D．威廉·约翰森

2．下列氨基酸中属于人体必需氨基酸的有哪些？（　　）

A．赖氨酸

B．色氨酸

C．亮氨酸

D．谷氨酸

3．让基因型不同而表现型相同成为可能的是下列哪项？（　　）

A．发育稳态

B．中性突变

C．沉默基因

D．各色环境

4．新性状起源的关键是什么？（　　）

A．新陈代谢

B．自我复制

C．自然选择

D．基因型网络

5．科学技术与自然创新的相似之处有哪些？（　　）

A．试错

B．人海战术

C．多起源

D．优化组合

6．达尔文进化论的局限在于无法解释什么？（　　）

A．遗传现象

B．新性状的起源

C．生命的起源

D．自然如何创新



想要获取生命多样性和进化动力的奥秘吗？
扫码获取“湛庐阅读”APP，
搜索“适者降临”查看测试题答案。

世界够大，时间够多

1904年春天，任职于加拿大麦吉尔大学（McGill University）、年仅32岁的新西兰物理学家欧内斯特·卢瑟福（Ernest Rutherford），在世界范围内成立了最早的科学组织伦敦皇家自然知识促进学会，并举办了一场演讲，演讲主题为“放射现象和地球年龄测定”。

当时的科学家对《圣经》里认为地球年龄只有6 000年历史的说法早就嗤之以鼻了，最广为认同的年龄测定是由另一位物理学家威廉·汤姆森（William Thomson）计算出来的，而他更为人熟知的称呼是“开尔文勋爵”。开尔文勋爵用热力学定律和地表导热系数测定出地球大约有2 000万年的历史。

从地质学的角度来看，2 000万年并不算很长，然而它的影响却非常深远。假如按照今天火山运动和地貌侵蚀的速度来算，2 000万年对于地球独特地貌景观的塑造根本不值一提，但达尔文提出的以自然选择为基础的进化理论却成了直接受害者。达尔文曾说，“威廉·汤姆森先生对地球年龄的测定极度困扰了我”，因为他知道地球生物从上一次冰河世纪结束后就没有发生过太大变化。据此达尔文推测，创造所有生物所经历的岁月必定非常悠久，不管是现存的还是已经成为化石的，2 000万年对于创造一个多样的生物界是远远不够的。

卢瑟福在发现放射性元素半衰期现象的几年后，逐渐发现开尔文勋爵的说法是错误的，他曾回忆道：“我走进演讲大厅，里面非常昏暗，但我还是在观众席中发现了开尔文勋爵，感觉甚是尴尬，特别是在讲最后一段关于地球年龄的部分时，因为我们的观点是相互冲突的。放射性元素会发生衰变并在衰变过程中释放巨大的能量，这类元素的发现使得我们对地球年龄的测算更加准确了。生物在地球上的起源时间得以大大提前，地理学家和生物学家提出的进化过程纵使缓慢，也成了可能。”

开尔文勋爵逝世于1907年，次年卢瑟福获得了诺贝尔奖。截至19世纪30年代，用放射性测量法估算出的地球年龄大约是45亿年，于是生物有了足够的时间在缓慢的进化过程中创造出多样性和复杂性，达尔文的进化理论也得以保全。

然而，真的是这样的吗？

作为自然界最出色的捕食者之一，游隼（falco peregrinus）是完美生物的代表。它有着极度轻盈的骨骼和健壮的肌肉，也是目前地球上飞行速度最快的动物。在旋转俯冲时，隼的飞翔速度可以超过每小时200千米。当隼俯冲而下用利爪抓住猎物时，由极高时速带来的冲击力几乎可以瞬间将猎物置于死地。即使不能，它也可以用锐利的喙折断猎物的脊柱。

游隼有一双锐利的眼睛帮助它们迅速捕捉到活动的猎物。隼的视力是人眼的5倍，也就是说就算在1.5千米之外，它们也能看清楚一只鸽子。和其他的捕食者一样，隼有一种所谓的瞬膜结构，又称“第三眼睑”。瞬膜有点像挡风玻璃，能在高速飞行时帮助阻挡飞尘并保持眼球湿润。隼的眼睛还拥有更多的光受体及视杆细胞，使它能在昏暗的光线下看清事物，甚至看到紫外线。

大自然的创造充满了神奇之处，但更奇妙的是，每一种不可思议的生物特性都是从一个极小的分子开始的，在漫长的世纪进化中，经过无数次的演变，最后交由大自然精挑细选。游隼的喙、爪子和羽毛的主要成分同人类的毛发和指甲一样，都是一种叫作角蛋白的蛋白质分子。色觉主要依靠视蛋白，而视觉的敏锐性与晶状体蛋白息息相关。

大约在5亿年前，世界上出现了第一种能够合成晶状体蛋白的脊椎动物，而视蛋白在7亿年前就已经出现了。它们出现的时候，生命已经在地球上居住了超过30亿年。对于那些复杂的生物大分子而言，30亿年的时间听起来就比较符合情理了。每个视蛋白和晶状体蛋白都是由20种氨基酸按一定顺序结合形成的多肽链，再由一条或一条以上的多肽链按照特定规则结合形成高分子化合物。如果只有一种氨基酸能够感知光波或是构成透明的晶状体，那我们要从多少条含有数百个氨基酸的多肽链中才能筛选出我们需要的那条呢？假设一条氨基酸链上有两个氨基酸，那么第一个氨基酸有20种选择，第二个也是，如此一来，总共将会有20^2种氨基酸链的可能组合。如果一条氨基酸链上有三个氨基酸，那么最终的组合方式将会是20^3种可能。以此类推，如果是4个氨基酸组成的多肽，将会有16万种氨基酸的可能排列方式。对于一条含有100个以上氨基酸的蛋白质（晶状体蛋白或视蛋白的链还要更长），可能的多肽链将超过10^{130}种。

为了让你感受一下这个庞大的数字，我们来举一个例子。宇宙中数量最多的原子是氢原子，物理学家估计氢原子的数量可达到10^{90}个，具体来说就是1后面跟着90个0。如此一来，蛋白质的可能数量就不仅仅是

宇宙级了，而是超宇宙级：远远超过了宇宙中的氢原子数目。要从这么多可能的蛋白质中找出唯一正确的那条，概率甚至比赢得自宇宙大爆炸以来的每一年的乐透都小。如果亿万种生物从生命出现伊始每分每秒都在寻找那条特定的氨基酸链，那么到现在为止也可能只尝试了10^{130}种蛋白质中极小的一部分，甚至都还没找到视蛋白。

17世纪的诗人安德鲁·马弗尔（Andrew Marvell）曾叹息，“只要我们的世界够大，时间够多”，为了避免那“无垠永恒的荒漠”，他无心关注时间长河下的自然奥秘，只想和情人享受片刻的欢愉。这里我们对这首诗的关注点在于悠远漫长的时间。在无尽的时光中，自然选择加上生物的变异，迟早会让一种叫游隼的生物进化出那双锐利的眼睛。达尔文进化论的主流观点是，优势性状赋予生物的优势，无论多么微不足道，都将在生物漫长的繁衍生息中被无限放大，这个观点解释了包括游隼在内的所有生物的多样性。

自然选择的神奇之处是毋庸置疑的，但它也有自身的局限性。自然选择能保留由变异产生的新性状，却不能创造它们。认为变异总是随机的观点，暴露了我们对变异的无知。自然界众多的生物性状，如果没有大自然对于生物进化的助益，其中许多近乎完美的结构可能永远都不会出现，而这种助益正是生物进化的能力。

过去15年中，在纽约和瑞士苏黎世大学的实验室里，在一群才华横溢的科学家的帮助下，我有幸得以对这些奇妙的自然规律展开研究。我们使用了在达尔文和卢瑟福年代无法想象的先进实验方法和精准的计算技术，希望解释生命多样性和进化能力的来源。目前的研究数据告诉我们，进化的奥秘远远不止我们的肉眼所见。神秘的自然规律隐藏在每个精巧的**DNA**里，隐藏在每个独一无二而又美丽动人的生命体中。

这本书的主题就是对那些自然规律的探索。

目 录

测试题　进化论的最大谜题：自然如何创新？

前言　世界够大，时间够多

01　达尔文进化论的局限

达尔文进化论的局限在于，它无法解释遗传现象。生命起源于何处？更好、更强的适者从何而来？大自然如何无中生有，如何创新？达尔文的进化论是人类历史上杰出的学术成就，但生物进化的秘密远不是达尔文进化论所能穷尽的。生物学在20世纪发生了翻天覆地的变化，现代技术得以带领我们探索生命进化的动力和起源。

02　新性状的起源

新性状的出现有赖于新的分子和合成这些新分子的化学反应的存在。生命以及生命背后驱动新性状出现的动力并不是神秘莫测的东西，这种动力本身和生命一样古老。我们还不知道生命到底是如何从最简单的形式进化出了如此高的复杂性，但我们知道，生命的开端不是一个自我复制的分子，而是一张新陈代谢的网络。

03　宇宙图书馆

一种生物所具有的全部生化反应构成了这种生物的新陈代谢。新陈代谢进化的本质在于重新组合。生命时刻在尝试每一种可能的基因新组合，重新解读，重新编译，然后重新布局代谢遗传，毫不停歇，从而造就并提升着代谢的多样性。新的代谢能力是不断驱动生命拓展最前沿阵地的引擎。

04　构型之美

蛋白质是生命的驱动者。每种蛋白质的构型都高度复杂，与它们执行的功能相适应。蛋白质的构型维持着生命世界的运转。大自然可以用蛋白质书写不同的文本，更多的文本就意味着更多的构型，参与更多的催化反应，执行更多的功能和完成更多的任务。

05　命令与操控

无论多复杂的生物，它的形态和功能都受到调节因子的控制。调节因子占据着某个基因相邻的一小段DNA，一旦它们遇上特定的DNA序列，就会与之结合。调节因子与相应的DNA需要在形态上互补。有些基因表达能被调节因子抑制，有些基因则需要它们激活。调节

因子指导着所有生物的发育。调节因子之间相互调控，形成了复杂的网络。

06　神秘的建筑师

多变的环境催生了生物的复杂性，而复杂性促成了发育稳态，发育稳态继而造就了基因型网络，后者让进化成为可能，使得生物能够通过演变适应环境的变化、提高自身的复杂性，循环往复，生物进化通过这种方式螺旋上升。这种进化方式的核心在于处在多维空间的基因型网络的自组织性。自组织性是生命绚烂光彩背后的支持者，它是隐藏的生命建筑师。

07　从大自然到工程技术

自然进化和技术创新有诸多共同之处，促进自然进化的基因型网络在人类技术进步中同样存在。与自然界类似，科研人员也总是行进在各自领域的最前线，他们依赖不断的试错、人海战术、多起源策略和组合优化，模仿自然的创造能力，实现技术突破和创新。技术发明的精简主义和高雅主义，深深隐藏在现实世界的背后。

后记　柏拉图的洞穴

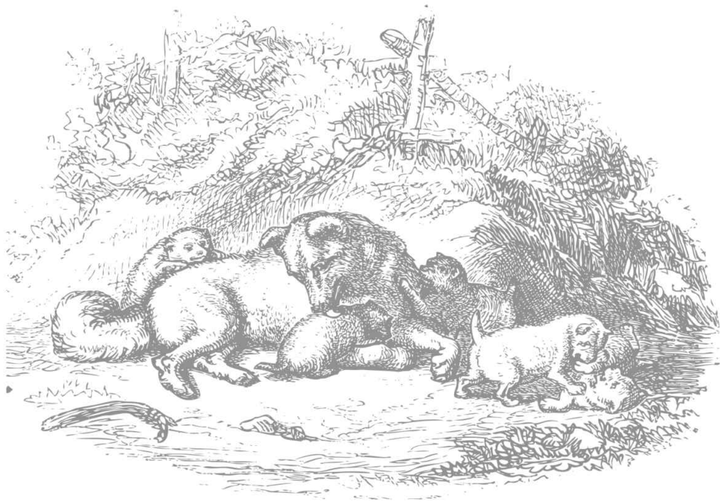
译者后记

ARRIVAL OF THE FITTEST

01 达尔文进化论的局限

Solving Evolution's Greatest Puzzle

达尔文进化论的局限在于，它无法解释遗传现象。生命起源于何处？更好、更强的适者从何而来？大自然如何无中生有，如何创新？达尔文的进化论是人类历史上杰出的学术成就，但生物进化的秘密远不是达尔文进化论所能穷尽的。生物学在 20 世纪发生了翻天覆地的变化，现代技术得以带领我们探索生命进化的动力和起源。



萨莉·加德纳（Sallie Gardner）可以算作世界上第一位电影明星。1878年，年仅6岁的“她”以惊艳的银幕处女秀宣告了电影的诞生。出生于英国的摄影师埃德沃德·迈布里奇（Eadweard Muybridge）想要解决一个当时让不少人都夜不能寐的问题：一匹奔马的四条腿会不会在某一刻全部离开地面？现在我们知道，答案是肯定的。而当时迈布里奇在马奔跑的路径上设置了24台摄像机，把一匹马飞奔而过的一系列照片用诡盘投影机[1]放映，萨莉就是那匹被拍摄的马。迈布里奇拍摄的布满噪点、镜头严重抖动的默片时长仅有一秒钟，这和21世纪初我们司空见惯的高清立体声环绕电影简直天差地别。然而从迈布里奇的片子发展到现代电影只用了近一个世纪的时间，并没有比达尔文发表的《物种起源》差多少。后者只比萨莉的亮相早了19年。

在那个世纪里，生物学领域的变迁甚至比电影技术更加剧烈。生物学革命打开了新世界的大门，如果是达尔文面对这些新图景，恐怕他的感受就像穴居人面对着浩瀚的宇宙。新的知识帮助我们解答了一个有关进化论的重要问题，一个达尔文和他之后的科学家都无法回答，甚至无法触及的问题：更好、更强的最适者从何而来？生命起源于何处？大自然如何能无中生有？

看到这里你可能不禁会疑惑，意识到生物可以进化并解释这种进化的发生原理，难道不正是达尔文进化论的伟大之处吗？不正是达尔文留给后人的财富吗？是，但也不是。毋庸置疑，达尔文的理论是那个时代乃至人类历史上杰出的学术成就。但生物进化的秘密远不止达尔文在进化论中所探讨的问题。事实上，达尔文甚至都没有意识到有关生物进化最核心的问题，更遑论解决。要说明来龙去脉，我们首先要看看达尔文在提出进化论的时候知道些什么、不知道些什么，他的进化论中又有哪些观点是走在时代前面的，而哪些不是。继而我们就会理解，为什么在一个多世纪之后的今天，我们才开始探讨“生命到底如何起源”这个问题。

人类早在达尔文生活的时代之前就已经开始关注生物的进化现象。2 500多年前，古希腊哲学家阿那克西曼德（Anaximander）——“日心说”的祖师爷[2]，认为人是由鱼变来的。14世纪的伊斯兰历史学家伊本·赫勒敦（Ibn Khaldun）则认为，生命会沿着从矿物到植物再到动物的顺序发生演变。许多年之后，19世纪的法国解剖学家艾蒂安·若弗瓦鲁·圣伊莱儿（Etienne Geoffroy Saint-Hilaire）根据爬行动物的化石总结

出，生物能够随着时间的推移发生变化。1850年，就在达尔文出版《物种起源》的9年前，维也纳植物学家弗朗兹·昂格尔（Franz Unger）提出，所有植物都是藻类的后代。另外，法国动物学家让-巴蒂斯特·拉马克（Jean-Baptiste Lamarck）则坚持，生物进化的动力来自“用进废退”[3]。

这些早期的学者似乎都预见到了生物进化的存在，然而，只要你稍微深究一下就会发现这些理论中的不实之处。比如阿那克西曼德认为人最初藏于鱼腹，待到孕育成熟，遂破鱼腹而出，诞于世间。这些与现今科学完全相悖的信条，在达尔文的时代依然大行其道。唯有一个观点受到了从古希腊到拉马克时代众多科学家的追捧：低等生物是由自然界的非生命物质自发生成的，比如湿泥巴。

在达尔文时代来临之前，进化理论已经拥有了众多支持者，当然反对的声浪也同样喧嚣。我所说的支持者和反对者与当今“年轻地球创造论”（young earth creationist）的信徒不是一回事，该理论的支持者普遍接受过半吊子的教育，往往自以为是、目空一切，他们相信地球是在公元前4004年10月的一个周六的夜晚被创造出来的。他们还相信诺亚方舟拯救了100多万种物种，只是诺亚可能忘了把恐龙带上船。鉴于当时诺亚已经600岁了，爱忘事似乎也情有可原。我所说的进化理论的反对者，都是当时科学界的巨擘，其中之一是著名法国地质学家、古生物学创始人乔治·居维叶（Georges Cuvier）。

古生物学的字面意思是“研究古代生物的科学”，例如恐龙。居维叶发现，古老岩层里的化石与年轻岩层中的差别巨大，而年轻岩层中的化石显示，它们与今天的生物十分相似。即便如此，他依旧坚信每种生物都是独一无二的，生物独特的形态不会变化，而只在极小的范围内存在个体差别。另一个反对者是卡尔·林奈（Carl Linnaeus），他仅仅比达尔文早出生了一个世纪。林奈是现代生物分类体系的鼻祖，然而这位分类学创始人直到晚年都视生物进化为谬论。

基督教的教义是解释这种抵触情绪最好的理由。对居维叶来说，他在化石中看到的生物多样性并不意味着生物可以进化，而是印证了造物主无与伦比的创造力。不过，还有一个更重要的原因则要追溯到古希腊哲学家柏拉图。柏拉图对现代西方思想的影响十分深远，20世纪的哲学家阿尔弗雷德·诺斯·怀特海（Alfred North Whitehead）曾直言，欧洲哲学的发展不过是循着“柏拉图的脚印”罢了。

柏拉图哲学深深植根于抽象的数学和几何学世界。在柏拉图的世界

观里，可见的物质世界反倒是海市蜃楼，不过是更高等的世界投射下的一掠缩影而已，那个更高等的世界是由各种图形组成的几何世界，比如三角形和圆形。对于柏拉图学派的人来说，篮球、网球和乒乓球有一个共同的本质，那就是球状的外形。每种球的物理特征无论如何变化，都不过是虚无的幻影，只有完美的、几何的、抽象的球形本质才是真实的。

对于像林奈和居维叶这样的科学家来说，要实现自己的目标，即把混乱无序的生物多样性以某种方式组织起来，柏拉图式的物种概念显得方便实用：每个物种都拥有区别于其他物种的不变本质。正是因为这种“不变的本质”，所以爬行动物中没有腿和眼睑的物种被称为“蛇”。在这种柏拉图式世界观的影响下，博物学家们的日常任务就变成了寻找物种的特质。这样说反倒是轻描淡写了，事实上，在本质主义的世界观里，“物种的特质”和“物种”这两个概念的界限是模糊的，特质即物种。

与之对比鲜明的恰恰是真实的世界，现实的自然界不断喷吐着新物种，并与原有的物种相互交融。生活在白垩纪晚期的真足蛇（eupodophis）拥有退化的后肢，而幸存至今的脆蛇蜥（glass lizard）则没有四肢。真足蛇和脆蛇蜥只是众多位于物种模糊边界的代表之一。生物进化的纷繁世界无疑是追求简洁和秩序的本质主义者的死敌。因此，当20世纪的动物学家厄恩斯特·迈尔（Ernst Mayr）称柏拉图以及他的信徒是“进化论者最伟大的敌人”时，也就情有可原了。

在帮助达尔文主义者占据上风的过程中，真足蛇化石只不过是证据之山上的一块鹅卵石而已。在达尔文生活的时期，分类学家已经将数千种生物归类，并且意识到了它们之间的相似性。地理学家已经发现地球的表面并不像看上去那样宁静祥和，新的地貌不断出现，板块之间时刻发生着折叠及岩层断裂。古生物学家在不同的岩石层中发现了不同年代的生命体，在较为年轻的地层里的生物化石往往和现今的生物相似，而那些在古老岩层里的化石则显得十分不同。胚胎学家已经向世人指出，在海里自由自在划水遨游的虾与偷偷附着在船体上远渡重洋的藤壶，在胚胎发育阶段十分相似。探险家，包括达尔文在内，则找到了许多发人深省的生物地理学模式。比如越小的岛屿上物种越少，同一个大陆东西两侧的海岸线上往往栖息着十分不同的动物种系，欧洲和南美洲的哺乳动物种类全然不同。

如果生物多样性建立在每一个物种被独立创造的基础上，那么局面就会像一团“剪不断，理还乱”的乱麻。而达尔文，有史以来最伟大的理

论学家之一，将它们编织成了自己理论中的美丽丝线。他无畏地向创世论者宣战，宣称所有的生物都有共同的祖先，把《创世记》从辩论桌上掀翻在地。

生物可以进化只是达尔文卓越的洞见之一，除此之外，他还提出了自然选择理论。这个自然界的中心法则是他在观察动植物选种的过程中偶然想到的。《物种起源》的整个第1章都在赞叹人类育种师培育的狗、鸽子、农作物以及观赏花卉的多样性。在短短100年里，人类就从同一个祖先中先后驯养出了大丹狗、灰狗、英国斗牛犬、吉娃娃等各种品类的狗。达尔文从这个令人惊叹的人工选择过程中意识到，自然选择应该也遵循着相似的原则，只不过它所历经的时间会更长、范围也更广。新物种的变异每时每刻都在发生，虽然绝大部分变异都稍显逊色，只有极少部分变异能够得到优等的性状。但无论优劣，它们都得符合一个相同的标准，那就是自然选择：只有适者才能得到生存和繁衍的机会。这个过程几乎完美地解释了生物多样性，遗传学家西奥多修斯·杜布赞斯基（Theodosius Dobzhansky）曾说："只有在进化论的光芒照耀下，生物学的一切才有意义。"

不过，这道进化论的光辉仅仅照亮了无数自然奥秘中的一小部分，还有一个它鞭长莫及的藏匿在黑暗中的疑问是：遗传机制。亲代将自己的遗传物质传给子代的时候，如果没有稳定的遗传机制作为保证，遗传性状，比如鸟的翅膀、长颈鹿的脖子、蛇的尖牙，就无法稳定延续下去。如果没有遗传，自然选择也就成了空中楼阁。达尔文对自己无法解释遗传的原因十分坦诚，他曾在《物种起源》中提到："遗传的法则仍旧充满未知。"这种真诚袒露自身无知的行为令人深感敬佩。



达尔文的理论就像萨莉奔跑的镜头，与静态摄影相比，那部时长一秒钟的默片在当时意味着革命性的超越，但离现代成熟的长篇电影依旧还有弱水之隔。事实上，在达尔文逝世的时候已经有人提出了遗传机制理论，只是人们并不知晓。在达尔文出版《物种起源》3年前的1856年，遗传学的奠基实验就已经开始进行了。令人唏嘘的是，即使是开展那个实验的科学家本人，也无缘在世期间一睹他的研究给生物学界带来的颠覆性改变。

这位科学家就是奥地利修道士格雷戈尔·孟德尔（Gregor

Mendel），他曾经就读于维也纳大学，之后便进入了布隆城圣托马斯修道院。在成为该修道院的院长前，孟德尔一直进行着豌豆实验，研究过的豌豆数量超过了两万粒。孟德尔在实验中特意选择了豌豆作为实验对象，因为豌豆有许多区别明显的相对性状：有的豌豆是黄色的，表面光滑；有的则是绿色，表面褶皱。最理想的是，这些性状都没有介于中间的过渡形态。类似的性状还有豌豆的花色、荚形和茎秆长度。孟德尔对性状不同的豌豆进行了杂交，并对大量的子代豌豆做了细致入微的分析。

孟德尔从对后代的研究结果中发现，同一个性状之间不会互相交融，比如第一代豌豆的表皮不是光滑就是褶皱，杂交得到的豌豆亦然，而没有出现介于两者之间的中间性状。另外，不同的性状以相互独立的方式遗传，杂交豌豆中黄色的豌豆可以是表皮光滑，也可以是表皮皱褶，而绿色的豌豆同样如此，因此子代的某些性状组合是第一代豌豆所没有的。每种遗传性状就像不可分割的基本单位，并且在遗传中呈现离散分布。从豌豆颜色和表面纹理的遗传中可以推测，豌豆总是成对携带控制每个性状的遗传单位，而在杂交时每个亲本只把其中一个传递给后代。只有这样，不同的性状才能以稳定而相对独立的方式进行遗传。

孟德尔在远离时代科学大潮的修道院里完成了他的研究，但他在最后犯了一个后来许多人都犯过的致命错误：他把自己的研究成果发表在了一本不入流的本地杂志上，那是一本以爱好自然为主题的刊物。更糟糕的是，在孟德尔逝世之后，他的继任者烧毁了他的著作。不过在孟德尔的论文发表34年后，“沉睡多年的睡美人”还是被荷兰植物学家雨果·德弗里斯（Hugo De Vries）唤醒了。德弗里斯独立完成了类似于孟德尔的实验。

时至今日，历史学家对于德弗里斯的研究究竟是自己独立完成的，还是剽窃了孟德尔的成果这一点依旧争论不休。毕竟，孟德尔的理论不仅姗姗来迟，而且迟了整整30多年，换谁都有可能希望借此让自己名垂青史。无论如何，德弗里斯唤醒了孟德尔定律，醒来的睡美人一发不可收，迅速在生物界确立了地位，成为一个全新的分支，也就是现在广为人知的遗传学。孟德尔式的遗传性状存在于许多动物、植物及人类身上。有些性状比较生僻，比如耳垢的黏稠程度（干或湿）；而有的性状则至关重要，比如血型种类（A型、B型或O型）；还有一些则与遗传病有关，比如镰刀形红细胞贫血病。

其实德弗里斯至少得到了一个慰问奖，他是遗传学名词“基

因”（gene）的提出者，这个词的重要性不言而喻。德弗里斯把孟德尔所说的遗传因子命名为“泛子”（pangenes），后来遗传学家威廉·卢德维格·约翰森（Wilhelm Ludvig Johannsen）又选择舍弃了前缀“pan”。

约翰森对现代生物学的贡献还包括另外两个重要的名词，他创造了“基因型”（genotype）和“表现型”（phenotype）这两个词，并对它们进行了定义。用今天的话来讲，基因型是指生物个体所有基因的遗传构成，而表现型则是生物个体表现出来的性状：生物的大小、颜色，是否有尾巴、羽毛或外壳等。从理解这两个词的区别开始，我们才能够进一步辨别生物进化中性状演变的因果关系。举例来说，生物学中有个词叫“变异”（mutation），200多年前人们就曾用它来表示生物体外观上发生的显著改变。

20世纪初期，变异既用于形容孟德尔式的遗传变化，同时也被用于表达单纯的外观变化，对生物体变化的因果关系研究造成了巨大的混淆。一个世纪之后我们才知道，变异改变的是基因型，比如远古动物体内视觉蛋白的变异。所谓的“变异”往往会影响生物的表现型，有些表现型对生物发育至关重要，比如只有视蛋白的出现，我们才能看到这个多姿多彩的世界。

只有辨清了基因型和表现型之后，我们才能探讨那个对理解生命进化无比重要的问题：变异到底是如何改变表现型的？这是达尔文没有解开的另一个谜题：新性状从何而来？新的变异，尤其是那些能够延长生物体寿命、增加异性吸引力、提高繁殖能力的变异到底从何而来？有人可能会觉得理所当然：变异和新性状的产生当然是随机的，听天由命。这种虚无的解释至今仍有不少拥护者，不过达尔文深知这个解释没有任何意义，他在《物种起源》中讨论变异的章节是这样开篇的：

> 一直以来，我自己都时不时把变异……发生的原因归因于天意。这种说法除了是彻头彻尾的错误之外，还暴露了我们对变异的原因一无所知的事实。

对达尔文来说，变异是个大问题，因为自然选择本身并不会导致变异。自然选择不创造新的变异体，而仅仅是对已存在的变异体进行选择。达尔文的确意识到了自然选择在生物进化中的正面作用，却始终无法参透变异的来源。

那么这个问题到底有多重要？试想一下，当今的我们和地球上最早

的生命体之间每一丝细小的差异，都意味着曾经发生过的一次进化，是生命面对生存的挑战时做出的适应性改变。这些挑战涉及方方面面，可能是把光能转化成化学能，或者把食物转化为能量，又或者是在栖息地之间长途迁徙。海洋里的每一汪水，陆地上的每一块草地、每一片森林和荒漠、每一个城市和乡村，地球表面的每个角落都存有生命的踪迹，每一个生命都在自己最适宜的环境中生龙活虎、繁衍生息，同时寻找着更优良的新性状。

这些适合生存的新性状，从最常见的光合作用、呼吸作用，到保护爬行动物的鳞片和为鸟类保温的羽毛，还有起到连接作用的结缔组织和内骨骼。有的性状相对复杂，而有的则相对简单。无论是小如仅有10微米的细菌鞭毛，还是大如3米长的蓝鲸尾鳍，它们存在的原因无非都是生命在进化中的某个阶段，出现了适应特定环境的新变异。

自然选择没有，也无从创造这些新性状。在达尔文去世几十年之后，雨果·德弗里斯清楚地意识到了这个问题：“自然选择可以解释最适者何以生存，却无法解释最适者如何降临。”如果我们无法理解最适者从何而来，那么我们也就无法解释当今生命所展示的惊人多样性。

生命具有进化的能力。不仅如此，生命在变异的同时依旧能够通过稳定的遗传保留已有的性状，它同时具有可变性和保守性。在20世纪早期，生物学家对其中的奥秘无从得知，这也在情理之中，因为离解决这些问题所需的生物实验技术和计算工具登场还有将近一个世纪的时间。

事实上，当我们回过头来看，20世纪早期的科学家意识到基因型和表现型的区别，就已经是一件非常了不起的事了。同孟德尔和迈布里奇一样，他们对自己所研究的东西充满了疑惑，甚至不确定“基因”到底是不是真实存在。它可能像重力一般无影无形，但也有可能切实存在，能够从生物体内分离出来并在实验室里单独进行研究。直到多年之后我们才知道，基因存在于染色体上，是由DNA构成的分子片段。

在发现基因的物理本质之前，先是由达尔文点燃了一场生物革命的星星之火，而孟德尔的发现则像一阵狂风使得火势肆无忌惮地蔓延开来。但是离散、单位化并不是所有遗传方式的特征，最简单的反例恰好来自我们的日常生活。比如，一个身高1.8米的男人和一个1.5米的女人生育后代，根据遗传的离散性规律，他们孩子的身高不是1.8米就是1.5米，不应该出现介于两者之间的中间值。但我们知道事实并非如此，他们孩子的身高在一个区间内呈连续分布。同样的道理，这些孩子的相貌、肤色、身形等亦然。达尔文之后的博物学家在自然界发现了许多呈

连续性分布的遗传性状：作物的产量、鸡蛋的重量、树叶的形状。总而言之，这种性状是大多数生物性状的遗传特征，它的重要性由此可见一斑。

离散和连续，到底哪一个对进化而言更重要？这一问题又激起了科学家们此起彼伏的争论。以达尔文为早期代表的自然主义者和渐进主义者倾向于关注微小的连续性变异；而另一些学者，如“孟德尔主义者”“变异论支持者”“突变论者”则倾向于关注孟德尔研究中的离散性突变。如果要给这个争论的双方拍一部卡通片，那么渐进主义者会说花园里的玫瑰是从它的某个五片花瓣的祖先一代一代进化而来的，而突变论者则会反驳说，只需要一次偶尔的“大突变”就能得到美丽的玫瑰，而无论它的祖先有多少片花瓣。

站在今天的角度来看，这个辩论跟中世纪学者们讨论得热火朝天的另一个问题不过是半斤八两：一个针头上究竟能够容下多少个天使跳舞？但是对于当时的达尔文主义者而言，这种辩论简直是噩梦。因为相比于自然选择，孟德尔主义者更相信突变在新性状产生过程中所起的主导作用。在他们眼里，突变才是生物进化的主要驱动力。德国动物学家理查德·戈尔德施密特（Richard Goldschmidt）曾把突变形容为“带来希望的怪物”，他举的例子则是为了适应海底生活而把双眼移到头顶的比目鱼。

虽然后来的研究证实孟德尔主义者的观点是错误的，大多数生物的进化的确有赖于漫长时间中自然选择的积累，但他们的观点也不是完全不对。困扰科学家多年的疑问不是自然选择，而是新性状到底起源于何处。但是孟德尔主义者关于变异的观点太超前了，在当时根本无法用科学的方法对遗传和变异给出解释，所以两大阵营的争论一直持续了整个20世纪。直到一个人们熟悉的观点再次进入大众视野，这场争论才慢慢平息并渐渐有了答案。这个观点就是：遗传和变异不仅仅发生在个体中，同时也是一种群体现象。



白色桦尺蠖（peppered moth）是一种不起眼的昆虫，白色的翅膀上散布着一些黑色斑点。在树干或者地衣上，黑白斑驳的翅膀是绝佳的伪装，不易被贪婪的捕食者发现。然而，如果某个控制翅膀颜色的基因发生了变异，就会导致黑色的桦尺蠖孵化，这些变异后的桦尺蠖无法有

效地伪装自己，因此很容易被鸟类发现。但是19世纪的工业革命却为黑色桦尺蠖助了一臂之力。那个时期的工业污染极其严重，树干和地衣都因为染上烟煤而变成了黑色，意外地成了黑色桦尺蠖的完美藏身之地，而白色桦尺蠖则不幸沦为捕食者的盘中餐。

如果自然选择当真起着重要作用，那么接下去会上演的一幕就是，随着时间的推移，黑色桦尺蠖会华丽逆袭，慢慢成为桦尺蠖群中的主流，而白色桦尺蠖将变得越来越稀少。这也正是19世纪在英国发生的事，黑色桦尺蠖的比例从1848年的2%猛增到1895年的95%。现象只是表面的，远没有它背后的实质来得重要：我们是否可以用某种方式预测优势性状在群体中的传播速度呢？或者相反，如果我们观测到某种性状在群体中的扩散速度，那么我们能由此推算出该性状的适应性是多少吗？这些通过数学进行量化的角度是原本的进化生物学不曾有过的，它导致了生物学领域一门新兴的独立学科的诞生：群体遗传学（population genetics）。

群体遗传学的核心不是研究某个生物个体，也不是整个种群的表现型，而是种群的基因池。举个例子，决定桦尺蠖翅膀颜色的基因有许多种，也叫等位基因，不同的等位基因决定着翅膀是白色还是黑色，它们在桦尺蠖群体中的分布比例和频率各不相同。

假设在某个时间点上，某个种群里两个等位基因的数量相同，但随后出现了一个新的影响因素，可能是一种新的天敌，也可能是环境污染，导致黑色桦尺蠖存活的时间更久，繁殖的后代更多。这个优势在最初或许并不明显，但哪怕对应黑色翅膀的等位基因只增加了微小的1%，从第一代中的50%增加到51%，那么随着时间的推移，这个比例就将持续增大，直到黑色变异体占据绝大多数，这就是自然选择：种群的等位基因频率在日积月累中影响着个体的性状比例。

这个观点具有划时代的意义。生物学研究的方式自亚里士多德以来就不曾发生过变化，生物学家总是先仔细观察，而后进行详细的实地或实验室调研，最后对观察结果进行详细记录，但是从群体遗传学开始，生物学家迷上了数学的力量，并把各种数学工具引入了生物学，包括微分方程和方差分析等。在各路科学巨匠，如休厄尔·赖特（Sewall Wright）、霍尔丹（J.B.S.Haldane）、统计学家费希尔（R.A.Fisher）等的共同努力下，群体遗传学能够相对精确地解决关于自然选择的量化问题。于是在同一时间，博物学家纷纷在野外研究桦尺蠖种群中等位基因的频率，而实验学家则在实验室里研究能快速繁殖的果蝇。数学像红

娘一样把原先井水不犯河水的两者一起牵引到了生物学的殿堂里。

群体遗传学中的新证据告诉我们，变异的概念极其宽泛，既有孟德尔式的离散性突变，也有连续性变异。孟德尔式的性状，如翅膀的颜色、豌豆的形状，都由等位基因中效力相对较强的主效基因控制；而连续性性状，比如身高，则是由多个微效基因控制的，每个基因都具有相同的效力。群体遗传学告诉我们，自然选择同时影响了这两种基因，但真正令人惊异的是自然选择在其中所起到的作用。

如果黑翅的等位基因降低了桦尺蠖被天敌捕食的概率，哪怕只是很小的几个百分点，它也能在经过几十代繁殖之后击败白色桦尺蠖而使黑色成为群体的主流。博物学家和实验学家都发现，微效基因的例子远多于主效基因，由此可见当年孟德尔在选择豌豆的时候有多么小心谨慎，毕竟他选出的性状都是由主效基因控制的，而这样的例子在自然界并不多见。进化在多数时候都是循序渐进的，不是一蹴而就的。

到了20世纪30年代，基于自然选择、遗传本质和种群思想的概念，诞生了一个新的理论：现代综合进化论（modern synthesis）。这个名字取自朱利安·赫胥黎（Julian Huxley）的同名著作。虽说是“现代”，但这个理论马上就有100年的历史了。和其他“百岁老人”不同的是，它没有任何衰老的迹象。在数学计算和数据分析的帮助下，这个理论更是稳扎稳打，获得了坚实的理论基础。现代综合进化论对人类生物学研究的各个领域，如追寻人类起源、研究人类迁徙、认识基因疾病等，都功不可没。如果这座知识殿堂有实体，那么几乎没有任何建筑能与它的华丽相媲美，无论是世界上最大的庙宇吴哥窟，还是艺术瑰宝泰姬陵，抑或是13世纪的哥特式大教堂。这是一座代表人类学术成就的宏伟殿堂。

然而，这个理论成功的背后同样隐藏着一个不太光彩的地方。现代综合进化论的创立者抛弃了生物体本身和表现型，一味执着于对基因型的研究。他们忽视了生物体本身的复杂和伟大性，有些生命体由上亿个细胞孕育而成，每一个细胞又由无数功能复杂的大分子组成。他们忽视了这些伟大的生命体是如何从一个简单的受精卵，经过无数精细而繁复的过程发育而来的，而基因又在这个过程中起了什么作用。

因为没有关注生命的复杂性，现代综合进化论的创立者侥幸避开了这个问题，结果是他们对进化最终的产物——生物体本身视若无物。为了能够把注意力全部放在基因型上，早期的现代综合进化论者将生物的表现型抽象为同一个概念：适合度（fitness）。适合度代表一个子代个体成功传递给下一代的平均基因数目，越是适应环境的生物对下一代基

因池的贡献也就越大。不仅如此，他们还假设每个基因对于个体适合度的贡献基本相当，例如，个体适合度是它的每个基因适合度的简单加和。

当然，我并没有批评的意思。现代综合进化论除了忽略生物整体之外几乎别无选择，因为用抽象的方式理解复杂事物总要付出代价：为了理解冰山的一角，你就必须用盲人摸象的方式忽略相对不重要的部分。当爱因斯坦说“事情应该力求简单，但是不能过于简单”时，天知道他到底想要表达什么。现代综合进化论的支持者只是在尽量简化这个问题而已，以便能够理解基因和基因型在进化中的作用。这个理论之所以能成功解释自然选择也正是因为摒弃了生物的复杂性。

但是当一个理论相对成功的时候，就很容易让人忽略它的局限性，这也是现代综合进化论在其鼎盛时期所犯的错误，生命的进化被重新定义，然后被贬低到了“基因库中等位基因变化”的层次。而最主要的局限性也使它无法回答《物种起源》中的第二个关键问题：新的性状到底从何而来？现代综合进化论解释了新性状如何在种群内传播，但还是无法解释它的起源。

当然，如果说所有进化论者都忽略了生物体本身，这样的言论未免有失偏颇，还是有一小部分进化论支持者在从胚胎发育的角度研究生物体的复杂性，但是这些胚胎学家却受到了现代综合进化论支持者的排挤。研究果蝇的遗传学家托马斯·亨特·摩尔根（Thomas Hunt Morgan）因为解释了基因与染色体的关系而在1933年获得诺贝尔奖。就在获奖的前一年，他说过这么一句话：“不管是用成年猿还是用猿的胚胎作为人类的祖先，其实真的无所谓。”

虽然群体遗传学家一直占据着生物学殿堂的前排座位，但那些在后排委曲求全的胚胎学家一直都没有放弃过希望，相反，他们一直在竭尽全力地向前排宣扬他们的主张。在20世纪后期，当进化发育生物学（简称“进化发生学”）开始作为一门新兴学科登上生物学舞台，誓要整合胚胎发展、进化学和遗传学的时候，那些胚胎学家曾经坚持不懈的呐喊声也渐渐得到了人们的关注。进化发生学对基因和胚胎的关系提出了全新的见解，解释了不同的基因如何像和谐的管弦交响乐团一样完美协作，从而使胚胎发育成为可能。

可惜迄今为止，还没有一个成型的理论能够和现代综合进化论相提并论。理论化是把散乱的事实修砌成一座学术大厦的唯一途径，而罪魁祸首正是我们上文中提到的生命的复杂性。直到今天，我们都要耗费九

牛二虎之力，才能勉强理解哪怕是最简单的生物体性状，前赴后继的生物学家孜孜不倦地研究了几十年也无从得知生物的基因到底是如何精确调控表现型的。如果说现代综合进化论者有一个牺牲了表现型而得出的遗传理论，那么胚胎学家手里则攥着众多生物的表现型，却没有任何可以拿出手的理论。

进化发生学告诉了我们一件很重要的事，为了理解生物新性状的产生，我们无法弃表现型于不顾。虽然我们无法全然了解一个生物体的复杂性，但是至少知道了某些表现型与生物进化的关系。这也是我们接下来的章节会继续探讨的问题。



前有达尔文，后有孟德尔，生物学在同一个世纪里发生了翻天覆地的变化，现代综合进化论又孕育了生物化学，一门在700多年前，从人类开始酿酒的过程中就初露锋芒的学科。酵母和糖是如何作用产生酒精的过程一直是个谜，直到达尔文发表《物种起源》的3年前，才由路易斯·巴斯德（Louis Pasteur）指出发酵是微生物作用的结果。短短几十年之后，巴斯德的结论就被推翻了。1897年，爱德华·比希纳（Eduard Buchner）证实，发酵的过程不一定需要生物参与，因为不含活体细胞的酵母提取物也能导致发酵。比希纳的发现加速了“活力论”的消亡，这个理论认为生命需要某种神秘的“生命力”，而生命力遵循着和非生命物体完全不同的自然法则。

比希纳除了告诉我们生命是基于化学的之外，更大的贡献是他发现了酶，这是一类由成百上千个氨基酸构成的巨大生物分子，它能加速化学反应过程。生物化学上一直沿用了比希纳的系统命名法为酶命名，即在酶的催化物后面加上“ase”的后缀。比如能水解蔗糖的酶就叫作“蔗糖酶”（sucrase），而能水解乳糖的酶叫作“乳糖酶”（lactase）。

比希纳的发现开启了生物化学领域一扇新的大门。他关注催化反应，而不是酶本身，揭开了化学世界的面纱，新陈代谢的过程也不再神秘莫测。广义来说，“新陈代谢”这个词来源于希腊语，原意是“改变”，主要包含两种类型。第一种改变是分解外源分子，比如葡萄糖分子，释放能量；第二种改变是生物体从外界环境中获取营养物质并转变成自身的组成成分，比如蛋白质中的氨基酸，同时储存能量。新陈代谢起着分解并排出代谢废物的作用。这些过程相对复杂，都需要酶的作用，涉及

上千个化学反应，从而使生物体能够完成能量交换和自我更新的过程。

蛋白酶对表现型的重要作用是20世纪一个具有里程碑意义的发现。同时它也为理解生物进化提供了新的视角：生物体无论发生多大的改变，都是从单个的蛋白质分子变化开始的。即便如此，它的光芒还是被另一个更重要的发现盖过了：基因的化学结构。

这一发现要追溯到达尔文1869年发表第五版《物种起源》的时候。瑞士化学家弗雷德里希·米歇尔（Friedrich Miescher）首先发现了一种区别于蛋白质的神秘物质，并称之为“nuklein”，但它的化学结构是几十年后才研究清楚的。直到1910年，这种物质被重命名为“脱氧核糖核酸”（DNA），包含4个碱基：腺嘌呤（adenine，缩写为A）、胸腺嘧啶（thymine，缩写为T）、胞嘧啶（cytosine，缩写为C）和鸟嘌呤（guanine，缩写为G）。1944年，奥斯瓦尔德·埃弗里（Oswald Avery）发现，将肺炎链球菌有毒株的DNA与无毒菌株混合，后者也会变得对老鼠有致死性。由此，生物学家意识到，DNA才是遗传物质的携带者。

在此之后不到10年的时间，詹姆斯·沃森（James Watson）[4]和弗朗西斯·克里克（Francis Crick）研究发现，DNA具有美丽的双螺旋结构，DNA双链像阶梯一样扭曲盘旋而上，每一个阶梯都由互补的核苷酸配对组成，DNA的碱基排列配对方式只能是腺嘌呤与胸腺嘧啶或胞嘧啶与鸟嘌呤。该结构也能顺利解释DNA的复制方式，进一步丰满了遗传的运作方式。至此，基因的定义已经远远超出了当年约翰森的想象。

从迈布里奇的诡盘投影机问世到彩色电视技术的诞生总共用了70年的时间，这是技术从在银器里记录黑白图像到用无线电把阴极射线管发出的电信号转变为光学图像所花费的时间。

在这70年间，生物学领域也发生了突飞猛进的变化，群体遗传学和现代综合进化论都在这个期间涌现，同时科学家还阐释了酶与DNA结构的奥秘（和彩色电视机的出现在同一时期）。化学知识在我们理解生物进化的过程中起到了无与伦比的重要作用，让我们离生命的终极奥秘又近了一些。

沃森和克里克的发现开启了分子生物学时代。在接下去的12年里，生物学家发现，DNA能够被转录为核糖核酸（RNA），随后在RNA转录为蛋白质的过程中，每3个碱基组成一个代表特定氨基酸的密码子（如图1-1）。3个碱基一组的密码子体系构成了64种不同的可能，大部分密

码子都与一种氨基酸对应，其中少数几个密码子比较特殊，它们与蛋白质翻译的起始和结束有关。

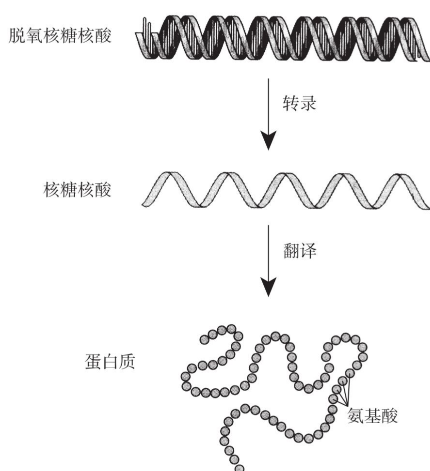


图1-1　转录-翻译

如果我们知道DNA的碱基序列，预测蛋白质链上的氨基酸序列应当是一件易如反掌的事。但事实上，蛋白质的结构不只是它的氨基酸序列那么简单，蛋白质盘绕成错综复杂的三维空间结构，要了解它们的功能，比如如何加速化学反应，我们必须知道蛋白质的结构和变化形式，然而至今我们都无法完全参透这个复杂的过程。从19世纪50年代开始，关于蛋白质如何折叠的研究就已经在血液的珠蛋白中展开，但是这些实验往往过程烦琐、耗时又长。通过DNA碱基序列预测氨基酸链不是什么难事，但是预测蛋白质的折叠方式就要复杂得多，就像要把爱尔兰诗人和剧作家叶芝的诗翻译成中文一样。

对于想要探索表现型来源的人们来说，这并不是什么好消息。想要了解生物体的表现型，不管是彩色的翅膀、敏锐的眼睛还是强健的骨骼，归根结底还是要了解组成生物体最基本的大分子结构。如果我们无法预测大分子的形态，就无法从基因型跨越到表现型。

不过每个蛋白质不都总是独立存在的，它们往往通过共同合作来应对机体复杂机制的作用，这让我们理解蛋白质的努力更是雪上加霜。以胰岛素为例，它是一种由胰腺分泌的，主要负责分解吸收葡萄糖的蛋白质分子，并能促进血糖进入肝脏。胰岛素无法直接进入肝脏，它是通过和肝脏细胞上的胰岛素受体相结合，受体会激活肝脏细胞内的另一些蛋白质，继而引发一系列连锁反应，促进葡萄糖分解的。我们的身体内每分每秒都在进行着类似的分子运动。自沃森和克里克发现双螺旋结构之

后，分子生物学家开始前赴后继地研究这一类问题。通过对一条条蛋白质链的研究，他们逐渐揭开了复杂大分子网络的神秘面纱，如那些控制人体感官和行为的大分子，甚至是任何一个方面的分子结构。

人类在这条研究之路上已经耕耘了很久，也收获了很多。走得越远，才越发现这条道路的漫长和蛋白质网络的复杂，从基因型转向表现型的探索也越加深远。

然而综观整个20世纪，仍然有很多支持进化论的生物学家完全不为表现型的复杂性所动。他们沐浴在现代综合进化论的阳光下，沉浸在对基因型的研究当中，这种执着在沃森和克里克的发现席卷了无知的人类之后，由于DNA分子序列识别新技术的出现而变得更加疯狂。这些技术也带动了一个新兴领域的诞生，叫作“分子进化生物学”（molecule evolutionary biology），主要研究氨基酸和DNA序列的变异。这项技术的前身就跟迈布里奇的诡盘投影机一样笨拙低效，一年时间只能研究不到几百个碱基对。而到了19世纪80年代中期，分析的效率提高了将近10倍，足以对人群中多个较短的DNA序列进行检测。

分子进化论者在这项技术的帮助下，发现了一件始料未及的事情：数量众多的基因变异在基因组中无处不在，甚至在那些数亿年中都没有发生明显改变的生物体内亦是如此。

分子进化领域一个早期的研究对象是醇脱氢酶，一种人体用于代谢酒精的酶。人类体内携带有这个酶的基因，果蝇亦然。我们不知道果蝇会不会因为啃食腐烂的水果而嗨得像摇滚乐队的歌迷一样，但我们至少知道果蝇对这些腐烂的水果趋之若鹜的同时，肯定需要醇脱氢酶来防止酒精中毒。1983年，哈佛大学的马丁·克雷特曼（Martin Kreitman）在一小群果蝇身上发现了这个基因的43种不同变异体。类似的变异也存在于人体当中，其中一种还会导致酒精过敏。酒精过敏曾在亚洲人的祖先中普遍存在，当时人们称之为“亚洲红脸症”（asian flush）。

但是克雷特曼在针对醇脱氢酶的研究中忽略了一个更大的秘密：大多数的基因变异是不表达的，它们改变了DNA序列，却没有改变醇脱氢酶的氨基酸序列。鉴于三核苷酸的密码子体系中，不少密码子对应的氨基酸相同，所以这种情况是可能的。但即便密码子具有冗余性，也不足以解释所有突变在遗传上表现出来的稳定性，毕竟突变有时候会穿插在密码子的3个碱基之间，从而彻底打乱遗传序列。所以，在突变中肯定还发生了一些不为人知的事。

这件事，就是自然选择。对酶分子不利的变异与它们对应的突变基因一起，早就在克雷特曼发现它们之前就被自然选择淘汰了。

克雷特曼的发现，以及其他类似的研究结果都反映了同一个现象：进化论思想中的进步与其他科学领域的改革不同。20世纪早期的量子物理学带来了和传统的经典物理学相冲突的世界观，而进化生物学的改革却丝毫不影响先前理论的核心观点。它们进一步深化、改造了历史，而不是推翻它。这些理论添加了层层的解释和方法，带来了新的视角。

正如电影《奔腾年代》（*Seabiscuit*）给萨莉·加德纳那第一次被记录下来的片子加上了颜色、音乐、对话和马蹄声，但是不会推翻迈布里奇对奔马四脚离地这一神奇现象的阐述。达尔文发现了自然选择的力量，现代综合进化论从基因频率的角度解释了自然选择，而分子进化生物学家则试图在DNA中寻找自然选择的蛛丝马迹，例如大量存在的不表达基因。不同的分支学科通力合作，渐渐揭开了达尔文留给世人的层层迷团。之所以不是所有的迷团，是因为分子进化生物学告诉我们更多的是有关生物基因的东西，而不是表现型，后者才是生物起源的核心问题。

克雷特曼在醇脱氢酶中发现的变异并不是巧合，类似的变异在自然界中广泛存在，甚至在活化石腔棘鱼中也有。人们曾经以为这种鱼早已灭绝，直到1939年又发现了幸存的个体。未表达突变的普遍性至今还在困扰着分子进化学家们：它们于表现型变化而言重要吗？它们和生物进化又是否有着紧密的联系？我们只知道，未表达突变的存在让基因型与表现型的关系变得更加扑朔迷离，表现型背后的原理依旧让人捉摸不透。



在20世纪80年代，光是掌握识别DNA碱基对的技术已经令人称奇。然而，与庞大的整个人类基因组相比，小小的碱基对就相形见绌了。人类基因组包含了30亿个碱基对，比《大英百科全书》还长10倍。我们身体内的每个细胞都拥有一套完整的基因组，高度压缩后形成了46条染色体。如大肠埃希氏菌这样微小的细菌都有450万对碱基对，比世界上最长的小说之一《战争与和平》的字数还多。高效测定单个个体的DNA序列所需要的技术还亟待改善，更不用说整个种群了。

发展这项技术的推动力来自“人类基因组计划”，这是于1990年启动的一个大型国际合作项目，由美国国家卫生研究院牵头。项目宗旨在于

了解导致疾病的基因，遗传病相当于一种特殊的新表现型。1998年，克雷格·文特尔（Craig Venter）[5]创立了塞莱拉基因科技（Celera Genomics）公司，立刻成了上述组织的强大竞争对手。文特尔的公司设法用1/10的成本测定所有的基因，并在2000年与公立组织在同一时间宣布完成了第一幅完整的人类基因组草图的绘制。

人类基因组是生物学领域众多的里程碑之一，它展示了无数的基因信息：人类所有的基因以及它们所编码的蛋白质序列等。克林顿总统在2000年的国情咨文中把人类基因组草图称为“生命的蓝图”。可惜的是，即使真如克林顿所言，那它也不过是一张陈旧的蓝图，我们无法从中搭建出它所描绘的宏伟景象，甚至都不知道该让建筑工人到哪里去施工。因为迄今为止，“人类基因组计划”依旧没有透露给我们任何与表现型相关的有用信息。许多人希望“人类基因组计划”能对关于一个人是否会得某种遗传疾病给出一个肯定的答案，而以下是克雷格·文特尔在2010年德国《明镜周刊》的专栏采访中关于预测基因疾病的陈述：

> 我们从基因组当中只能得出遗传疾病发生的可能性而已。在临床医学中，如果告诉你罹患某种遗传病的可能性是1%或3%又有什么意义呢？这些信息一文不值。

这个评价虽然苍凉，却是不争的事实。其中的理由或许你已经猜到了：基因型和表现型的关系复杂得难以想象。雄心勃勃犹如“人类基因组计划”，也只不过是从基因型出发，前往表现型途中的又一个一公里而已，这条路的尽头依旧遥不可及。

虽然“人类基因组计划”有它的局限性，但也带来了许多益处，其中一个就是DNA测序技术的蓬勃发展。2000年，一个操作者能在24小时内读取完100万个碱基对；到了2008年，测序仪器已经能够在相同的时间内测定10亿个碱基对。这项技术还在迅猛发展着。在你阅读这两行字的时间段里，基因组测序的成本就可能已经从1 000美元降到几美分了。这些技术使得研究人类和其他物种的基因变异成为可能，它们把种群基因学上升到了种群基因组学的高度。

种群基因组学的诞生意味着基因型研究的终点，但对表现型来说却并非如此。在20世纪50年代中期，有关蛋白质的功能以及相互作用的研究就已经启动，科学家们一路高歌猛进，势如破竹。但时至20世纪90年代，他们就不得不转换研究思路了。以胰岛素为例，先前的研究已经让

我们明确了合成胰岛素所需的基因，以及这些基因所编码的蛋白质和功能。但这些信息无外乎“谁是谁”或者“谁知道谁”，它们只是对信息进行了明确和组合，而对于预测个体的表现型，例如一个人是不是会得糖尿病，则丝毫没有用处。

科学家努力得到的结果还不足以告诉我们关键的细节，例如一个过程中涉及的蛋白质分子数量为多少，或者分子之间的关系强弱为几何。糖尿病的病因涉及几十种蛋白质大分子，每一种对糖尿病的患病都只有几个百分比的助益，它们之间通过相互作用对诱发糖尿病产生微妙的影响。所以单纯系统地罗列所有相关的蛋白质分子以及它们各自的特性，对于我们理解生命过程而言收效甚微。我们需要弄清楚不同分子之间是如何相互协作的。

处理这种整体性的唯一手段是数学，数学能够消化大量的实验数据，从而描述生物大分子的活动和密度是如何随时间变化的，这些活动是理解表现型的关键。举个例子，II型糖尿病发病时身体会发生胰岛素抵抗，这是一种与健康人完全不同的表现型：胰腺释放胰岛素，但由于肝脏对胰岛素不敏感，所以从胰岛素受体开始，激素信号会在传递的某个环节突然减弱或增强。这个改变影响了信号链，因而诱发了疾病。只有数学的精确量化能够帮助我们理解这种微妙的过程，这是单纯的罗列和分类做不到的。

然而，用数学方法描述表现型并非易事，从数十年的实验数据来看，主要大分子相互之间的相互作用有许多变量。这些计算的复杂性绝非简单的人工笔算所能完成，即使是最杰出的数学家也做不到，必须要有计算机的协助。

21世纪生物学对计算机的依赖性，犹如摄影技术之于相机。计算机的适用范围绝非仅限于实验室，从超低温冰箱到咖啡机，它们凭借自身强大的能力在各个领域占有一席之地。就像17世纪的显微镜一样，计算机带领我们走进了一个新世界，一个如此微小的世界，即使是最尖端的电子显微镜也无法欣赏得到，即分子的世界。称计算机为“21世纪的显微镜”当之无愧，可以帮助我们看到连达尔文都不了解的分子网络。

生物学领域中，计算机技术的整合是一个新兴现象。纵观生物学的发展历史可以看到，生物学的发展总是受制于数据处理能力。早期探险家需要航行数年，才能在偏远的小岛上发现新的物种；即便在分子生物学发展早期，分离一个基因也通常需要花费好多年时间。如今这种景象已经一去不复返了。由于科学技术的发展日新月异，生物信息数据如雨

后春笋般喷薄而出，你不仅可以在数千个不同的数据库中找到基因和基因组的信息，还能找到许多其他生物大分子，以及这些大分子之间的相互作用关系。每年都有大量的新数据进入数据库。新一代的科学家——计算机生物学家，只负责处理现成的数据即可，而无须自己进入实验室收集信息。生物学家摇身一变成为信息科学家，享有着无穷无尽的数据信息。在探讨自然法则的过程中，限制我们的仅仅是自己的想象力和分析数据的技巧。

当然，这些技术也会面临相应的挑战，因为生物性状起源的问题已经困扰了科学家将近一个世纪的时间。一方面，我们知道生物的表现型就像一幅巨大的点彩画，作画的人每次只往画上加一点。但是，这个比喻并不能告诉我们具体应当如何创作出一幅美丽的图画。研究性状起源的挑战很容易让人望而却步。以醇脱氢酶为例，它的氨基酸连接方式已经远远超过宇宙中的氢原子数。如果我们用完全的随机突变来解释新性状的起源，那么这首从达尔文时期就开始回荡的咒歌与阿那克西曼德的鱼腹理论似乎半斤八两，不啻于把我们的无知藏在地毯下假装看不见。当然，这并不意味着突变和自然选择就不重要。不过仅有自然选择不足以解释自然界惊人的有序性，我们仍然缺少一种能够加快进化速度的方法。

哪怕时间倒退几年，我们都不可能理解这种方式，更不要提这本书的出版。由于生命体由分子构成，所以我们需要通过分子来了解进化：不仅是DNA中的基因，还有基因型究竟如何塑造了表现型。表现型和DNA本身并不对等，它是生物体有序的层级架构，从最高层的器官到组织，再到细胞，再往下还有构成细胞的分子和分子之间形成的关系网络，最后精确到单个蛋白质。新的表现型和性状可以在这之中的任何一个层级出现。30年前，我们对于这种复杂性还一无所知。

如果连如今的我们都只是略懂皮毛，那就更不用提达尔文了。把他不知道的东西列出来简直可以出一本现代生物的百科全书。达尔文不但不了解生物性状的起源，在前孟德尔时期，他对基因的存在同样茫然无知，更不用说DNA和遗传密码了。他同样也不会知道群体遗传学和发育生物学，他对分子如何构成生物体一无所知。达尔文对生命真正的复杂性毫无察觉，许多后人也因此觉得他们可以理直气壮地忽略这一点。但是为了找寻生命进化的秘密，我们必须勇敢面对生命的复杂性，而不是逃避。

一种久经考验的认识生命复杂性的方法是关注一个或几个基因型以

及它们对应的表现型，这也是早期基因学家发现基因的基本方式：通过某个表现型的变化追溯源头的变异基因。在基因组时代，这个方法也适用于研究DNA序列的功能：诱变某个基因并观察相应的表现型变化。应用不同技术得到的发现相当惊人，比如苍蝇体内的基因突变导致它发育出了两对翅膀，植物长出了变形的叶子和以新物质为食的微生物等。科学家诱变了许多基因，得到了千奇百怪的表现型。

然而，这些个别的例子到底能在多大程度上说明问题呢？就像探险家如果要绘制新大陆的地图，光是沿着海岸线航行，随便抛锚上岸散个步是远远不够的。他们需要环绕整个大陆以画出它的轮廓，从河流三角洲驶入内陆摸索清楚河流的分布，他们还必须爬上山脊，穿过沙漠和丛林。对于生命的创造性，我们也需要绘制这么一张地图，一张从基因型到表现型的地图，标出每一个基因型的变化，以及它们如何影响了表现型。我们需要这样的地图来补全达尔文的伟业。

不过即使拥有最好的技术，这张地图也没有那么容易绘制。就一张具有高分辨率的地图而言，我们需要获得超过10^{130}种氨基酸链的表现型资料，那还不算由成百上千种基因和蛋白质组成的更高层次结构。换句话说，绘制一张高分辨率的生命地图不只是困难，几乎是件不可能的事。幸运的是，我们并不需要把每一粒沙子都在地图上描绘出来，如果我们只关注地形特征，就能减轻很多绘制的负担，需要研究的基因型数量也会大大下降，不过剩余的基因型数量依旧数以亿万计。鉴于表现型可研究的角度很多，所以我们要精心选择，保证这些我们研究的角度对生命的进化而言至关重要，同时又处于现有知识和分析工具所能处理的范围之内。

柏拉图的本质主义论与进化主义论不共戴天数十年之后，在这些地图中正东山再起。与柏拉图时期简单枯燥的几何世界相比，21世纪本质主义的内涵要丰富得多。它对达尔文主义思想兼容并蓄，又不拘一格，是我们理解自然选择的关键。仅凭肉眼人类是无法了解某些现象的，就像无法用肉眼看清楚萨莉·加德纳在奔跑的时候是否真的四脚离地。幸运的是，我们现在已经具备了看清进化世界的技术。

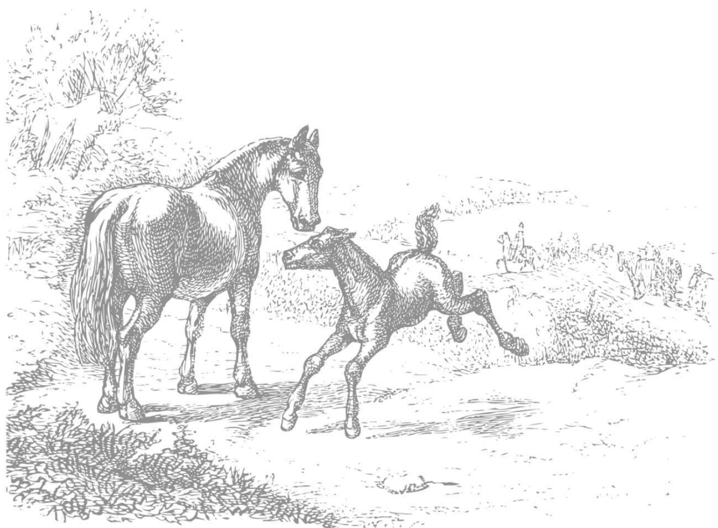
现代技术给我们展示了一个柏拉图式的色彩斑斓的世界，展示了40亿年以来生命进化的动力和起源。

ARRIVAL OF THE FITTEST

02 新性状的起源

Solving Evolution's Greatest Puzzle

新性状的出现有赖于新的分子和合成这些新分子的化学反应的存在。生命以及生命背后驱动新性状出现的动力并不是神秘莫测的东西，这种动力本身和生命一样古老。我们还不知道生命到底是如何从最简单的形式进化出了如此高的复杂性，但我们知道，生命的开端不是一个自我复制的分子，而是一张新陈代谢的网络。



有一个你在家里就可以尝试的实验。找一个容器，往里面装一些小麦，拿旧的内衣裤封住容器口，然后等上大概20多天，你就会发现容器里出现了老鼠，有新生的幼鼠，也有长大的成鼠。这个现象是由17世纪的医生及化学家扬·巴普蒂斯塔·范·海尔蒙特（Jan Baptista van Helmont）首先发现的。他还发现，在阳光照射下，两块砖之间的罗勒叶能够生出蝎子来。

范·海尔蒙特并不是“自然发生说”的首创者，这个学说的起源至少可以追溯到亚里士多德，然而他的确是这个学说最后的拥护者之一。时至今日，任何声称小麦和旧内衣裤相互作用之后能够产生生物的科学家都会被打上“妄想狂”的烙印，但是在范·海尔蒙特的时代，他粗糙的实验和结论却没有为他招来坏名声，相反，范·海尔蒙特于1644年在人们的敬仰中逝世。在“自然发生说”被广泛接受的年代，人们认为范·海尔蒙特的实验只不过是证明了一个显而易见的事实罢了。

范·海尔蒙特逝世数年之后，一名来自意大利的医生弗朗切斯科·雷迪（Francesco Redi）才向世人展示了这个实验的正确做法。在广口瓶里放上肉块，不消一会儿肉上就会爬满蛆虫，但这些蛆不是由肉块自发产生的：雷迪用一块棉布盖住广口瓶之后，由于苍蝇不能在肉块上产卵，蛆也就没有再出现。

雷迪的工作加速了“自然发生说”的消亡。在这方面同样功不可没的还有17世纪的荷兰纺织品商人兼镜片打磨师安东尼·范·列文虎克（Antonie van Leeuwenhoek），他发明的显微镜打开了通向微生物世界的大门。曾几何时，由于超出肉眼可见的范围，未知的微生物世界成为“自然发生说”拥护者最后的庇护所。直到18世纪中期，依旧有人认为腐烂的有机质可以产生微生物，这种观点的拥护者不乏像苏格兰牧师约翰·尼达姆（John Needham）这样的社会名流。一个世纪之后，路易斯·巴斯德才证实尼达姆本末倒置了：是微生物引起了有机质的腐烂，而不是腐烂的有机质孕育了微生物。巴斯德通过对肉汤以及肉汤附近的空气进行灭菌处理而使其免于腐烂，给“自然发生说”的棺材钉上了下葬前的最后一颗钉子。

虽然巴斯德证明了生命的自然发生并不存在，但他以及同时代的其他人都不知道生命究竟起源于何处。在当时，生命起源的问题属于化学研究的范畴，而不是生物学。而19世纪的化学家与那些在20世纪初苦苦

思索变异来源的孟德尔主义者，都面临着一个同样的问题：他们生得太早了。那是一个德米特里·伊万诺维奇（Dmitri Mendeleev）还没有发明出化学元素周期表的时代，对生命的化学元素展开研究更是一片大大的空白。由于起源于声名狼藉的炼金术，现代化学经历了漫长的岁月才成为一门受人尊敬的科学。即便如此，哪怕已经进入20世纪，当诺贝尔奖获得者、著名量子物理学家沃尔夫冈·泡利（Wolfgang Pauli）的妻子与一名化学家私奔之后，泡利在一封给朋友的信中仍写道："她哪怕和一名斗牛士私奔也好啊，可是她却偏偏选择了一个平凡的化学家……"

在过去的一个世纪里，我们知道了生命体复杂多样的表现型正是"自然发生说"面临的最大困境。如果一个拥有特定氨基酸序列的蛋白质分子都不能自发形成，那一个包含了数百万种蛋白质和其他复杂分子的大肠杆菌又怎么可能凭空出现呢？现代生物化学使得我们能够估算一个大肠杆菌自然发生的概率，在此前提下，复杂生命体自然发生的概率几乎为零。

不过这并不意味着自然发生在生命出现的早期阶段没有出现过。事实上，早期生命的出现甚至需要自然发生的帮助，只是其产物的复杂程度远远比不上现代的细胞及蛋白质。打个比方，地球上最早的生命就好比牛车上的一个轮子。这个轮子也是经由步步打磨而成，并非一蹴而就。虽然漫漫历史的泥潭已然抹去了这些步骤的痕迹，不过化学家还是觅得了些许蛛丝马迹，这也正是我们本章的主题。化学家不仅说明了早期生命出现的过程，还证实了一个更重要的假说：今天自然界所有生化反应所遵循的原则，与生命出现之前的无异。无论古今，新性状和最适者的出现都需要新的化学反应过程和分子作为前提。



冥古宙（Hadean Eon），指40多亿年前的地球，也是地质历史的开端。冥古宙的名字取自古希腊神话中的地底世界，早期的地球也的确是一个地狱般的地方。地球诞生之初表面覆盖着炽热的岩浆，大气里弥漫着炙热的浪潮。即便后来表面的岩浆冷却，凝结成坚硬的地壳，地球也不是什么生机勃勃的地方。如果有天外来客拜访过冥古宙时期的地球，它们将看到地球表面遍布无数坑坑洼洼的火山，还有滚烫的蒸汽雨落进原始海洋里。如果不是巨大的大气压（当时大气的密度远远大于现代大气），地球上的海洋早就蒸发干了。

无须多说，人的身体根本无法承受在这样的大气里呼吸的压力，更别提大气里有致死剂量的二氧化碳和氢气。另外，抱头找掩体也是个好主意，因为在一个叫晚期重轰炸期（Late Heavy Bombardment）的阶段，许多巨大的小行星不间断地撞击着原始的地球。如今，地球上大部分小行星的撞击痕迹已经被地质运动抹平，但你依然可以在夜空的月亮上看见巨大、阴森的环形山。通过岩石中含有的化学时钟，即随着时间推移稳定衰减的放射性元素，比如铀元素，我们就可以推算出这些小行星，乃至地球的年龄。

地球早期历史中最惊人的莫过于最恶劣的时期过去之后，生命出现的速度，这开始于大约38亿年前。在之后的大约4亿年间，地球上出现了迄今为止发现的化石证据中最古老的微生物。西格陵兰岛（West Greenland）岩层中的碳氢同位素指示出，在38亿年前左右，最古老的新陈代谢反应已经出现。这意味着生命是利用时间的一把好手，在应当登场的时候毫无延误地出现在了地球上。这样看来，生命以及生命背后驱动新性状出现的动力似乎并不是多么神秘莫测的东西。驱使进化发生的动力本身和生命一样古老。

地球上生命的起源需要用化学理论来解释，其中最早的理论被称为“原始汤”假说，人们通常认为这个理论的提出者是亚历山大·奥帕林（Alexander Oparin）以及霍尔丹，正是那个在20世纪20年代提出现代综合进化论的霍尔丹。值得一提的是，富有洞见力的达尔文早在他们之前半个世纪就有过类似的想法。在1871年写给朋友约瑟夫·道尔顿·胡克（Joseph Dalton Hooker）的信中，达尔文推测说：“如果有这样一个温暖的小池子（这个如果是多么异想天开啊），里面有各种氨磷盐，另外还有光源、热源和电等，这里的蛋白质能够自动形成继而参与到更复杂的后续反应中。”与此同时，达尔文也告诉了我们为什么如今找不到这种“温暖的小池子”：以如今生物的代谢速度，池子里的成分会立马被现今的生物体吸收以至吞噬殆尽。

“原始汤”理论一直作为一个假说存在了数十年。直到1952年，诺贝尔奖得主哈罗德·尤里（Harold Urey）位于芝加哥大学实验室的研究生斯坦利·米勒（Stanley Miller），为这个假说提供了强有力的证据支持。在合理推测地球早期大气的主要成分之后，米勒把这些气体密封在一个容器内，以电火花模拟原始大气中的闪电，并用冷凝水模拟降雨。几天过后，许多有机分子——那些通常由生命体合成的成分，出现在了米勒的迷你世界里。这个实验的意义非凡，因为它显示在我们居住的这颗行

星动荡不安的年轻岁月里，有机质能够从无机质转变而来。米勒的原始海洋里出现的有机质并不是普通的有机质，而是组成现代蛋白质的原料分子：氨基酸，如甘氨酸和丙氨酸。后续的实验中甚至出现了许多其他组成生物体的物质，包括糖类和DNA的组分物质。米勒实验最重要的意义在于，它把有关生命起源的讨论从哲学思考上升到了实验科学的范畴。

1969年9月，人们在比1952年米勒模拟的原始地球更恶劣的环境里发现了生命物质。那年9月的某一天，默奇森（Murchison）上空突然出现了一个爆炸的火球，这个有着数百号居民、位于墨尔本北部约160公里的小镇上空犹如出现了第二个太阳。在爆炸发生之后，陨石在空中留下一道浓烟后碎裂成大大小小的碎片，最大的一块落在了一座谷仓里，所幸没有造成伤亡。这起陨石坠落发生在人类首次登月两个月之后，当时的科学家对于任何研究天外来石的机会都心痒难耐。

在默奇森陨石中，当时的科学家发现了不得了的东西。在来到地球之前，默奇森陨石已经在太空里游荡了数十亿年，它的年龄几乎和地球一样。就是在这块陨石里，科学家发现了数种构成蛋白质的氨基酸，以及作为DNA主要成分的嘌呤和嘧啶。在后续研究中，应用21世纪的分光镜技术进行的分析显示，默奇森陨石中含有超过一万种不同的有机成分。

默奇森陨石并不是自然界的一朵奇葩，我们有必要知道这一点：在无数其他来自宇宙的陨石中也同样发现有有机物质的存在。幸运的是，宇宙中的分子由于结构不同而存在不同的辐射发射与吸收特征，技术发展到今天，借助极度灵敏的射电望远镜，我们已经不需要等到陨石撞击地球，就能根据波长信号区分出遥远星云中数百种有机成分在不同波段的喃喃细语。实际上，应该说它们简直是在“呐喊”。星云物质中3/4的成分是有机物质，其中就包括类似甘氨酸这样组成生命的关键成分。让人意外的是，星云中数量最多的三原子分子是水分子，这不得不让我们重新思考地球是不是真如我们一直以为的那样特殊。

组成生命的成分在宇宙中十分常见，不禁让人联想到地球上的生命可能来自宇宙。陨石和彗星，尤其是在地球形成之初撞击地球的那些，它们带来的水是现今地球海洋水总量的10倍，带来的气体则是现在大气总量的1 000倍。不仅如此，它们还带来了星云中丰富的有机分子，这些有机分子起到了至关重要的作用。很可能有10万亿吨的有机碳，甚至百倍于此，从外太空进入了地球。那相当于当今在生物圈中流通碳元素

总量的10倍。扫过地球公转轨道的彗星尾部尘埃尤为重要，因为不像需要经历着陆时高温爆炸的陨石，其中许多有机成分会遭到破坏，彗星的尘埃会温和而持续地向地球播撒生命的种子，润物无声。倘若如此，也许宇宙尘埃才是我们真正的母亲。

生命的成分到底是来自外太空还是诞生于地球，也许我们永远都无从得知。不过，从天文观测中我们还是能得到许多简单而重要的启示。首先，只要环境条件合适，组成生命的物质成分是可以自然发生的。其次，所谓合适的环境并不像达尔文描述的“小池子”那样近在咫尺而又得天独厚。它可以远在数光年之外，也可以像星云那样在宇宙里随处可见。

还有一点是关于直到今天依旧适用的新性状的：新性状的出现有赖于新的分子和合成这些新分子的化学反应的存在。为了理解新性状出现的原理，我们有必要先探讨生命物质分子的起源。



组成生命的物质分子并不是生命本身，就像一堆砖头和木材根本算不上是一栋大楼。至少，生命还需要一张包含许多获取能量、合成生物体所需物质分子的化学反应网络，这张网络也被称为新陈代谢。生命还需要有增加自身数量的能力，即自我复制，以遗传的方式将自己的优势特征传递给子代个体。如果没有子代对亲代性状的遗传，达尔文主义者的进化论就成了空谈，自然选择也就没有了意义。

不过这并不意味着新陈代谢和自我复制总是两者兼有。即使在你生活的周围，这两者也不总是同时存在的。病毒可以自我复制但其本身并没有新陈代谢的能力，它们通过劫持宿主细胞作为自身新陈代谢的厂房。真正的生命体必须同时拥有新陈代谢和自我复制的能力，而这导致了我们遇到的第一个“鸡与蛋”的问题：到底是先有新陈代谢，还是先有自我复制？

也许是出于对DNA双螺旋结构的喜爱，主流科学界曾经一度认为是自我复制首先登上了历史舞台。但是由于现今自然界存在的自我复制现象非常精致而复杂，要解释它的起源可不是件轻而易举的事。此外，DNA的脱氧核苷酸序列只是遗传信息的载体，它们不能自我复制。DNA会首先被转录为RNA，RNA再被翻译为相应的蛋白质（如图1-1），而蛋白质才是生物功能的执行者，包括转录和复制。

生物体的功能都由拥有不同氨基酸序列的蛋白质合力完成，没有一种蛋白质可以单独完成这些任务。如此精确复杂的分工又引发了另一个“鸡与蛋”式的问题，这次是有关于蛋白质和核酸（核酸是RNA和DNA的总称）的，这两者到底是谁先出现的呢？考虑到我们之前所说的概率问题，要求两者在自然界同时出现似乎有点不切实际。如果最初的生命是以一个自我复制体的形式存在的，那么这个“亚当”或者“夏娃”分子必须足够有能耐才行，它既要能储存自身的遗传信息，又要能自我复制。

当1953年发现双螺旋模型的时候，沃森和克里克就已经意识到，DNA复制的关键在于DNA碱基对的互补性：鸟嘌呤与胞嘧啶配对，腺嘌呤与胸腺嘧啶配对，这种配对将双螺旋的两条单链黏着在一起。他们的原话是，这“马上就让人联想到了一种遗传物质复制的可能机制”。这种机制几乎当即就把蛋白质作为最早复制体的可能性排除在外，没有像DNA双螺旋分子那样将两条单链配对的简单互补原则，由氨基酸组成的蛋白质无法以沃森和克里克所说的方式传递遗传信息。

综上所述，蛋白质并不是一种理想的自我复制分子。但是核酸似乎也没有比别的分子好到哪里去。核酸能够胜任蛋白质执行的生物功能吗？它们能够催化自身的复制吗？甚至，它们真的有催化活性吗？DNA分子的作用和结构似乎注定了这些问题的答案都是否定的。DNA最基本的任务是储存信息，为此它可以牺牲其他一切。它懒惰、保守，在生物体中一代又一代地保持传递。所以在酶被发现之后的半个多世纪里，科学家一度认为只有蛋白质可以催化化学反应，而核酸则没有这个能耐。

这让第一个自我复制体究竟是何方神圣显得扑朔迷离。直到1982年，化学家托马斯·切赫（Thomas Cech）和西德尼·奥尔特曼（Sidney Altman）才把RNA从丑小鸭变成了白天鹅。RNA曾经一度是分子生物学的继子，备受冷落。它的主要作用是将DNA的遗传信息转移到核糖体，后者是细胞内一台庞大而复杂的蛋白质合成机器。但前述两位科学家却发现，RNA能够在某些化学反应中起到催化剂的效应。

RNA也能像蛋白质一样催化化学反应的惊人发现，本身就像一剂科学的催化剂。很快，科学家们就意识到RNA拥有久远的历史，甚至比蛋白质和DNA都要古老，在生命混沌初开的时候，RNA就是那个失落世界里的统治者。不过，和失落的亚特兰蒂斯不同，早期的生命世界还是为我们留下了许多它存在过的线索。RNA曾经作为生命体关键分子的证据之一，便是它当今仍然在生物体中所起的核心作用。举例来说，核糖体由数十种蛋白质以及数种RNA分子构成，而在装配氨基酸、合成蛋白质

的时候起到催化作用的恰恰是那几种RNA分子，而非蛋白质。事实上，这些蛋白质本身恰恰是通过RNA催化合成的。

远古时期，RNA可能同时肩负着储存遗传信息和催化自我复制两种作用，但我们对于它如何做到这点却一直百思不得其解。为了说明最早出现的生命形式，我们不妨将起源之初的生命抽象为一个能够自我复制的简单分子。这个单分子将非常类似于RNA复制酶（RNA replicase），一种能够催化RNA复制的酶。

如今，世界上一些最优秀的化学家正在全力寻找这种简单的复制酶。他们迄今为止最好的成果是合成了一段长度为189个核苷酸的RNA，这段RNA具有一定的增殖行为，但它远不具备自我复制的能力，能够作为模板进行复制的区域仅包含其中的大约14个核苷酸。但是这依然启发我们，如果能够解决几个关键问题，RNA自身催化复制是完全可能的。其中一个主要的问题恰恰在于碱基互补性。

互补的碱基对会自动配对，也就是说一条母链和互补的子链能够退火[6]成一条双链RNA，就像双链DNA的形成过程一样。为了复制出更多的RNA，双链分子必须要解旋为单链，以便每条链上的信息可以被阅读。不过一旦你或复制酶将双链分开，互补的碱基对就会马上退火，像透明胶一样互相黏着在一起。所以对于RNA的自我复制而言，成也碱基互补，败也碱基互补，这是一把双刃剑。

最初的复制酶面临的另一个问题是必须绝对精确，因为任何复制错误都会导致误差灾变（error catastrophe）[7]。这个模型最初是由诺贝尔奖得主、化学家曼弗雷德·艾根（Manfred Eigen）发现和提出的。

如果要理解误差灾变模型，不妨想象一下中世纪抄写宗教经典的僧侣，他们逐字逐句地抄写枯燥的经文，如果有一个僧侣抄错了一个单词，那么这个错误的单词就会被另一个僧侣继续抄录下去。同样的道理，其他僧侣也可能在抄录经文的同时混入自己的错误，一传十，十传百，日复一日，年复一年，总有一天，宗教的经典会变成一堆逻辑混乱、毫无意义的文字垃圾。

RNA复制酶也面临着同样的问题，它如同一本分子经文，只是在RNA的世界里有一点小小的不同：复制酶本身既是经文，又是抄写经文的僧侣。RNA复制酶是一本自我抄写的书，抄录过程中出现的错误不仅影响文本本身，还会同时影响它本身复制的能力。这就好比犯错的僧侣不光写错了经文，他所犯的错误还让后来抄录经文的继承者头脑不清，

变得更加容易犯错。

只有那些几乎不犯错的复制酶才能保全核酸酶本身的遗传序列，从而保全其自我复制的能力。如果复制酶的准确性太低，催化产物多数为有瑕疵的复制酶，效率低下，或者催化复制更加不准确，随着时间的推移，这些催化产物最终会降解为无用的分子碎片，最初的编码信息也随之丢失。1971年，在曼弗雷德·艾根获得诺贝尔奖4年之后，他尝试计算了如果要规避误差灾变，复制酶应当具备的复制准确度。计算结果显示，复制酶的长度越长，所需的精确度就越高。套用一个简单的估计方式，一个长度为50的复制酶需要低于1/50的复制错误率，而长度为100的复制酶的错误率则需要低于1/100，以此类推。即使是我们在上文中提到过的那个长度为189的“最佳成果”，它的复制错误率依旧数倍于此。即使它可以完整地复制自己，所得的分子后代的命运也不过是径直滚下误差灾变的万丈悬崖，万劫不复。

幸运的是，生命在这方面的造诣远远超过当下的人类。催化DNA复制的蛋白酶，其误读率低于$1/10^6$。这种精确性的代价是其作用方式的高度复杂性。催化复制的酶包括一些功能高度专精的蛋白质，它们负责校对和修正其他酶的复制错误，这相当于有一群分工明确的僧侣，互相检查抄录的经文内容。编码这些蛋白质需要相当长的基因，远非原始的RNA复制酶可以相比。为了确保遗传信息复制的完成度，RNA复制酶催化的复制反应必须高度精确。你或许会发现一个新的“鸡与蛋”式的问题已经呼之欲出了，它的另一个名字是艾根悖论（Eigen's paradox）：精确的复制需要庞大而复杂的酶分子进行催化，而庞大和复杂的酶分子则需要精确的复制来保证。直到今天，大自然也没有为我们指出任何解决这个悖论的出路，不过我们将会在第6章中看到，生物的进化方式为我们提供了些许线索。

互补的RNA分子之间顽固的黏着性，以及要命的艾根悖论，都让“自我复制先于新陈代谢出现”的观点显得岌岌可危。但是如果和接下来的第三个问题相比，它们简直就是珠穆朗玛峰山脚和山顶的区别：从哪里获得充足的原料以满足复制的需要？复制所需的原料是富含化学能的分子，它们包含了几乎所有需要的化学元素，包括碳元素、氮元素以及氢元素。举个例子，现代生物体中的蛋白质催化DNA复制时，每秒钟需要消耗大约1 000个脱氧核苷酸分子。

即便最初出现的复制酶效率非常低下，每秒钟只能消耗一个脱氧核苷酸分子，大概需要三分钟才能完成自身的复制，由此可以看出，复制

对于原料的需求依旧不会因此而降低。原因在于，每个复制酶在一变二之后就分别成为一个复制酶，酶的数量以及这些酶催化生成更多酶的能力也随之增加。以现代的眼光来看，虽然早期复制酶的催化效率奇低无比，但是以一变二的复制方式依旧导致了指数级的增长方式和对复制原料的巨大需求。只需6个小时，这种增殖方式就需要消耗掉1吨核苷酸，一天之内消耗掉2.5吨，而一周后这个数字将超过80万吨。

生命的本质，正是一支贪得无厌、如狼似虎吞噬高能物料的分子大军，和所有行军的队伍一样，一旦切断补给，生命就会迅速崩溃。不仅如此，鉴于达尔文进化论和自然选择建立在物种大量繁殖，即复制的基础上，如果没有持续供应的食物链，两者都将成为空谈。另外，复制酶也和士兵一样争强好胜。在竞争中处于下风的分子最终将会由于复制不出足够数量的本体遭到淘汰，而饥饿会加快劣势分子消失的速度。没有足够的原料，生命就如同一根受潮的火柴，在贫瘠的地球上昙花一现，而后销声匿迹。

米勒的实验以及外太空播撒到地球的化学物质，都不足以支持早期地球上的那支饥饿的军队。虽然它们都带来了生命的重要组分，比如氨基酸，但是仅凭它们还远不足以解决早期生命的温饱。米勒当初的实验花费了数日才由1 000克无机碳获得几微克的有机分子。而纵观整个地球历史，虽然从古到今陨石为地球带来了数以百万吨的有机碳，但远水救不了近火。在地质史早期，嗷嗷待哺的复制酶未必能够等到从天而降的那一块陨石。指望陨石养活地球上最早的生命，相当于你坐在家里并期待每隔几天，就有一辆运送肥料的卡车撞进自家后院的花园里。

虽然双螺旋充满美感的结构诱惑着它的拥护者们竭力维护“自我复制早于新陈代谢出现”的观点，但上述三个问题还是不由得让人怀疑这是本末倒置。自我复制优先理论的支持者们幻想出了一家光鲜亮丽的汽车工厂，却忽略了零件供应商的重要性。没有零件供应商为工厂提供足够数量的轮胎、轮轴、变速器以及引擎，工厂里再高通量的流水线也不过是形同虚设，毫无意义。如果供应商效率低下、货源不足，导致工厂几年才能生产出一辆车，那么产量缩水、工厂倒闭就几乎不可避免。这个困境的解决方法显而易见：在第一个能够自我复制的分子出现之前，一张为生命提供各种原料的化学反应网络就已经准备就绪，为生物体源源不断地提供所需的物质。

换句话说，生命的开端不应当是一个可以自我复制的分子，而是一张新陈代谢的网络。



伴随恰当的分子出现，为生命提供能量和所需物质的化学反应最后也应运而生，但是这个“最后”并没有那么轻描淡写，生命的出现经历了相当长的时间。如果没有外界的帮助，生物体内的某些化学反应需要数千年才能完成。因此，新陈代谢需要催化剂，生物体内的催化分子可以显著提高反应的速度。催化剂的一个突出特征是：它们的催化效应与热力学有关。热是原子和分子的无序运动的结果，催化剂会改变反应分子之间的碰撞和接触，同时自身在反应中保持不变。催化剂在新陈代谢反应中煽风点火，它的主要作用是降低一个特定化学反应所需的活化能，从而成倍地提高反应的速率。现代新陈代谢中化学反应的催化剂几乎全部为酶，它们是极其高效和复杂的蛋白质分子，一种酶严格对应一种化学反应，某些酶还能将所催化反应的速度提高万亿倍。我们的身体里有数千种不同的酶，失去任何一种都可能让我们像得不到食物补给的原始复制体一样崩溃。

但是，38亿年前还远没有蛋白质催化剂这么先进的好东西。达尔文可没有提到他的“小池子”里有酶，这也是为什么许多科学家不再追捧他的池子理论的原因。另一个问题在于，分子如果要发生反应就必须先发生接触。由于分子在水环境中进行着热力学的无序运动，所以分子发生接触是一个随机的概率事件，概率的大小与给定环境中的分子密度成正比：分子越少，发生的反应就越少。也就是说，如果没有分子的高度集中，新陈代谢也就无法发生。如果早期海洋里的原始生命浓度过于稀薄，生命也将难以为继。这也是为什么化学家需要在试管里而不是游泳池里做实验的原因。如果被冲进茫茫的原始海洋里，新生的化学分子将一去不复返。

有人提出了潮汐池模型，以弥补达尔文的“小池子”本身的不足。在这个模型里，低潮期水因吸收热量蒸发而导致池中的化学物质浓缩，汛期涌入的水则起着搅拌的作用。但是和早期地球上恶劣的环境相比，这种模型里的水池简直犹如度假地的海水浴场。地球形成之初，月球公转轨道的半径只有现在的1/3，月球掠过地球上空时猛烈地拖拽着地球上的海平面，掀起的浪头高度是如今的至少30倍。此外，月球围绕地球公转的周期大约为5个小时，也就是说每隔几个小时它就会在地球上引起汹涌的浪头，根本没有给生命成分留下浓缩的机会。

在过去的几十年里，进化生物学向着更精致、更小的试管实验不断

求索，苦苦追寻却一无所获的科学家意外在深海中找到了一些答案。1977年，潜水调查船“阿尔文”号在加拉帕戈斯群岛（Galápagos Island）附近超过2 000米深处的太平洋海底发现了一个世外桃源。科学家发现那里到处都是两米长、长着红色羽毛样饰物的无嘴管虫，生着腿、用金属矿物武装贝壳的螺类，还有眼睛退化的虾类。海床上铺着厚厚的由微生物组成的菌毯，这些科学家从没见过的微生物同时也是海底其他生物的食物来源。与这些怪异的生物本身相比，海底生态圈维持自身运作的方式则显得更加匪夷所思：生态圈所需的物质补给直接来自地球母亲，那些从地壳裂口喷涌而出的炽热的营养物质、化学能量以及达尔文的“小池子”所缺乏的催化剂，造就了海底生态圈的繁荣景象。

低温海水穿过炽热的裂口，下沉到岩浆房附近而被加热到沸点。而后沸水又上升，就像大气中受热上升的空气，直到它与上方的低温海水相遇、混合为止。在穿越海底火山的旅程中，海水穿过地壳并从中滤走大量的矿物质、气体和其他营养物质。当海水降温时，这些物质就如同空气中的水汽凝结成雪花一般沉淀下来。和雪花不同的是，这些沉淀的物质日积月累，在海底形成巨大的“烟囱”，高度甚至能超过60米。在“烟囱”生长的同时，它们还会不断喷吐出水和沉淀的颗粒物，看起来就像真的烟一样。“烟”的颜色或白或黑，主要取决于其中的化学成分。

显然，从裂口里升腾出的热水是海底生命的能量来源，但热量并不是最重要的。熬出生命“浓汤”的不是热量，而是热水中丰富的物料成分。裂口中的海水里含有丰富的化学物质，例如作为臭鸡蛋气味来源的硫化氢。海底火山的这些成分对我们来说是纯粹的毒药，但是对海底某些种类的细菌来说则是肥沃的养料。与植物吸收光能并利用二氧化碳合成复杂分子的光合作用不同，海底细菌能够进行化学合成。它们可以利用无机分子、海底火山中丰富的碳元素以及其他化学元素合成自身所需的有机成分。化学合成也不是海底生态中唯一存在的自养方式。虽然海平面以下2 000米的地方一片漆黑，几乎没有任何光能够穿透到那里，但海底火山依然散发着微弱的火光，足以让某些细菌利用这些光能进行合成反应。虽然海底火山生态圈供给生命的方式非常怪异，但这种方式非常有效，从而使得这些围绕火山存在的世外桃源有着千倍于周遭贫瘠海床的繁盛。

如果说达尔文的“小池子”是一碗平静温和的浓汤，那么深海高温的火山就是一口粗暴原始的高压锅。火山口里的海水受到一段高达1 000米的水柱施压才没有在高温下沸腾，水柱的压力高达约200个大气压，

几乎相当于每平方米200吨质量所产生的压强。作为现今地球表面最高温度纪录的保持者，海底火山如此极端的环境依然没有能够阻止生命诞生，着实令人惊异。一种名为甲烷火菌属（methanopyrus kandleri）的细菌能够在超过122摄氏度的环境里繁殖，这已经超过了微生物学家用来给实验设备进行灭菌的温度。甲烷火菌属在温度达到130摄氏度的时候依然能够存活，但是会停止繁殖。

自从达尔文乘坐贝格尔号造访之后，加拉帕戈斯群岛已然成为一个研究进化生物学独特而富饶的户外实验室。这座火山群岛上有着巨大的海龟、独一无二的海鬣蜥以及调皮的加拉帕戈斯海狮。后来，人们在距离加拉帕戈斯群岛大约400公里的地方发现了另一个独特的秘境，它就是海底的高温火山。实际上，这种人们从前闻所未闻的生态圈其实随处可见，由于地心熔岩从海床裂开的缝隙中喷涌出来，数以千计的火山群在全世界海洋的底部喷吐着滚滚浓烟。众多的海底火山相连形成一条巨大的火山链，形成中央海岭，直通地球深处。从这条延绵纵横的裂口里流出的岩浆不断改变着地球表面的地貌。

与网球上的缝合线类似，这条海岭环绕着整个地球，周长相当于落基山脉、安第斯山脉以及喜马拉雅山脉总长度之和的4倍多，超过地球周长的两倍，而它全长都隐藏在海水之下。与它的长度同样惊人的还有海岭中火山链的滤水量：每年有超过200立方千米的海水穿过炽热的火山口，这意味着每过10万年整个海洋的水就会在中央海岭完成一次循环。

海底的高温火山口已经成为生命发源地的热门候选，但并不完全是因为火山周围发现的坚韧而又原始的各色生物，更重要的原因在于海底火山周围富饶的海水中，蕴藏着丰富的能量和化学物质。另外，这些火山已经历经岁月，几乎和海洋本身一样古老。早在生命出现之前它们就已经开始喷吐营养物质。时至今日，海洋中的水已经经过海底火山数万次的过滤，足以将生命的种子播撒到世界各地。

不止如此，海底火山还解决了几个一直以来困扰水池模型的问题。海底火山为生命的诞生提供了大量试管环境，由于从冷却的高温海水中析出的矿物晶体形状复杂多变，由这些晶体堆叠而成的海底“烟囱”在结构上充满气孔和通道，每一个孔道都相当于一根迷你试管，显微级尺寸的分子得以在这里接触并发生反应，而不会被冲进茫茫大海。你完全可以把这些海底“烟囱”想象成塞满数百万个反应试管，同时还在不断壮大的实验室。

如果规模还不足以解决所有问题的话，这些实验室还备有催化剂。这里说的催化剂并不是酶，而是诸如铁硫化物、锌硫化物之类的矿物，它们要么以颗粒形式悬浮在海水里，要么覆盖在孔道表面。除了催化剂，高温水和低温水的混合还带来了额外的好处。高温会同时加速合成以及降解生命成分的反应，火山口中心炽热的高温让生物的成分分子变得不稳定，而火山口周围冰冷的海水又会导致生命反应极度缓慢。正是由于火山口海水混合形成的水温梯度，保证了原始生命化学反应所需的最适宜温度。

深海热泉也很可能是研究新陈代谢起源的最佳场所。但是即便我们有十足的把握这么猜测，研究生命起源的科学家对于新陈代谢是否起源于深海也还是不置可否。因为我们实在无从考究到底是哪一个新陈代谢反应最早出现在生命的历史上。最合理的猜测可能是那些存在时间最久远的代谢反应，是那些不管是人类、动物，还是植物和微生物都拥有的代谢反应，甚至包括海底火山附近那些坚韧的微生物。在所有符合这个条件的可能代谢反应中，有一个显得尤其醒目：一个名为三羧酸循环[8]的循环反应。

三羧酸循环包括以柠檬酸为起始分子的10步反应，柠檬酸得名于它让柠檬具有的口感。三羧酸循环反应在经历众多步骤，生成名字相当生涩的众多中间产物，诸如丙酮酸、草酰乙酸、乙酸等之后，最终以生成两分子的柠檬酸而结束。

一个分子通过循环反应变成两个分子，这听起来让人觉得难以置信。不禁让人联想起19世纪声名狼藉的永动机。不过这个反应实际上没有违反任何物理定律。三羧酸循环的本质是一个柠檬酸盐分子分解为两个小分子，然后利用二氧化碳中的碳元素以及其他分子的化学能，逐步合成新的柠檬酸分子。

科学家在地球上最古老的生命体内发现了三羧酸循环中的部分反应，但三羧酸循环出现在我们古老的祖先体内，并不是科学家猜测它是最早出现的新陈代谢反应的唯一依据。三羧酸循环的许多中间产物是合成许多其他生命必需物质的原料：草酰乙酸为许多氨基酸以及脱氧核苷酸的合成提供原子团，丙酮酸为糖类的合成提供原子团，乙酸则是合成脂肪的原料，所有这些都是细胞膜的重要组分。当然，它们也是许多其他生物分子的原料。如果你要寻找一个新陈代谢的中心反应，三羧酸循环是当仁不让的选择。

三羧酸循环还是一个重要的可逆反应，它可以朝正向或反向进行。

其中一个方向，也就是上文所述的反应，就像无机电池驱动引擎制造生命所需的原料。栖息在海底热泉附近的细菌赖以为生的化学合成，正是利用了这个原理。如果这个反应逆向进行，就可以为维持生命活动的电池供能，我们的身体正是利用这个过程从食物中获取化学能。

即便三羧酸循环有着古老的历史，它的中间产物是合成反应的枢纽，并且是一个双向都具有重要意义的可逆反应，我们仍然需要一个米勒那样的实验作为它的证明。遗憾的是，世界上暂时还没有这样的实验。由于海底热泉的环境非常极端，在实验室里进行模拟的难度远远超过米勒的实验。此外，海底“烟囱”的反应孔道结构复杂，表面还包裹了无机催化剂，这两者对早期生命的出现至关重要，这样的试管可不是轻易就能在市场上买到的。虽然我们还不知道整条循环反应如何自然出现，但已有科学家指出了一种可能的方式：在铁硫化物或锌硫化物这类催化剂的催化下，三羧酸循环的关键分子丙酮酸首先在高温高压的环境里出现。在丙酮酸的基础上，实验室里自发出现了循环中剩余的其他反应。

三羧酸循环还有一个诱人的特征：循环反应的结果是分子数量的增加。每一次循环结束，初始的一个分子就变为两个，新生成的两个分子各自开始新的循环，而后生成四个分子，以此类推。化学家把这种现象称为自催化反应（autocatalysis），这也是从最原始的RNA复制酶到现代细胞生命的决定性特征：它们都在不断地复制自己。

但是三羧酸循环的自催化与RNA复制酶自我复制的本质不同。和循环里的其他中间分子一样，柠檬酸并不是直接复制它自己，而是通过完成整个循环中的反应，间接进行复制。我们假想的RNA复制酶是一种可以自我增殖的分子，相比之下，柠檬酸只是一张自催化反应网络的产物。这不能说是三羧酸循环的缺点，相反，它给我们的启示是，RNA复制酶以及它所拥有的遗传信息可能并非生命的决定性特征。换句话说，遗传可能出现在生命诞生之后。

我们目前不知道，也许不久以后可以弄明白，三羧酸循环是不是所有新陈代谢反应的鼻祖。我们也不知道是不是在RNA复制酶之前真的有新陈代谢反应出现。不过确切无疑的是，地球历史上第一个能被叫作活物的东西，不论它是什么玩意儿，都需要自催化反应来解决自己的温饱问题。生命所需的新陈代谢可不是区区几个反应，因为每一个反应都需要许多其他代谢反应提供原料，以保证充足的代谢物质。一旦工厂和供应商都就位，达尔文的进化论就开始展现威力了。进化论使得相对优秀

的工厂保留下来，与这些工厂相关的、更出色的供应商也就得以保全，后者又反过来造就了更优秀的工厂，以此类推，在无尽的循环里支撑起所有的生命之舟。

鉴于科学家发现的另一种罕见的催化剂，上述循环反应能够在深海热泉里诞生可能并非完全出于偶然。蒙脱石（montmorillonite），得名于法国的一个小镇蒙脱城（Montmorillon），当地农民利用这种黏土矿石在盐碱旱地里储存水源。20世纪末期，吉姆·费里斯（Jim Ferris）等化学家发现了蒙脱石的一个新作用，它可以让组成RNA的小分子自动装配成超过50个核苷酸长度的RNA链。



当新陈代谢和自我复制准备就绪，生命就几乎要从一片混沌之中涅槃而出了。不过它还缺一身合适的行头，现代所有的生命体都在用相同的材料包裹自己：两亲性（amphiphilic）的脂质分子。“amphiphilic”的词根来自古希腊语中的“both”（双）和“love”（亲）。由于一端含有亲水基团，而另一端含有疏水基团，就像水坑里的一滴油会在表面散开一样，两亲性的分子同时“亲”水和“亲”脂。

如果你有机会观察一下两亲性的脂质在水里的表现，肯定会大吃一惊：脂质分子能够自动形成囊泡。这是一些由一层薄薄的膜围成的空心球体，脂质分子在膜上的排布方式如图2-1所示。乍一看，我们可能很难理解这些分子要如何在没有外界的安排和帮助下，自动排列成如此复杂和有序的结构，但事实上并不难：这种排列是同时符合分子两端基团亲和性的最佳方式。图中实心圆代表的亲水部分距离水最近，而疏水部分离水最远，两层脂质分子相互为疏水基团起到隔绝水环境的作用。当你往水中加入脂质分子，这种膜就可以自发生成。此外，它们还在以自催化的方式生长，囊泡体积越大，生长得就越快。

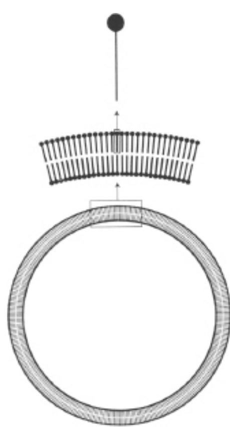


图2-1　生物膜

囊泡膜成分的起源并不神秘，也不遥不可及。三羧酸循环里就有脂质分子的前体产物，另外，像默奇森陨石那样的地外来石也是这类分子的重要来源。你可以用热水浸泡陨石粉末的方式制造出这些自动装配的囊泡。不仅如此，催化RNA成链反应的蒙脱石，同样可以加速脂质膜的自动装配。深海热泉环境的帮助还远不止于此，它还能浓缩膜成分。这个发现来自哈佛大学的杰克·舒斯塔克（Jack　Szostak）实验室，他们模仿构建了海底热泉中的孔道结构并发现，在极其微小的毛细管中，加热后的脂质分子浓缩并聚集到了同一侧，随后开始形成囊泡，而这一切都是自发的。

只要成分正确，复杂的结构就能凭空出现，这让人多少嗅到了范·海尔蒙特“自然发生说”的味道。不过，两者存在着本质的区别。老鼠、蛆虫或细菌的自然发生，需要借助无法解释的神秘或超自然力量，比如活力[9]。在活力论面前，由比希纳发现的酶显得滑稽而可笑。相比之下，生物膜和生物分子的自发装配，或者说是自组织（self-organization）形式，只需要简单的物理学和化学常识就可以理解。膜结构的装配只需要大量相似分子之间的相互吸引，就像海底火山喷发的颗粒自发堆积成高耸的海底“烟囱”，或者在蒙脱石催化下延伸的RNA链。以自组织形式形成的膜和分子在自然界算不上是什么稀罕的玩意儿。

自组织在宇宙中随处可见，甚至平常得往往会被我们忽略。自组织的出现远早于生命以及自然选择，它是恒星和星系出现的原因，也是地球诞生的推手，地球继而通过自组织俘获了月球，获得了海洋和大气，这股洪荒之力还在持续改变着板块的位置。自组织造就了小到显微镜下的雪花的对称结构，大到狂怒的台风云，另外还有沙丘变幻的轮廓以及

晶体永恒的美丽形状。如果说生命的起源中同样包含了自组织，我们也不用感到惊奇，因为自组织的确无处不在。

生命的自组织生物膜模型能够解决另一个有关早期生命的谜题：第一个细胞进行分裂的方式。现代细胞分裂的方式极其精致和复杂：由数十种蛋白质通力合作挤压并分开细胞，同时确保每一个子细胞都获得一份完整的母细胞DNA拷贝。脂质囊泡的分裂则显得相对原始和简单，舒斯塔克的团队在2009年观察到了快速生长的脂质囊泡在分裂过程中的性状改变，即球形的液滴在分裂时逐渐变为细长的空心管。这些空心管非常不稳定，轻微的碰触就会让它们破碎成一个个小的液滴。更神奇的是，当研究者把RNA分子置入空心管时，它们会被分配到后来形成的小液滴里。没有生命的脂质液滴能够像细胞一样分裂：只需要借助体系内各成分简单的化学特性，而无须借助活力论，并且完全是自发的。

虽然我们已经从最开始的原始汤理论一路走到了这里，但是面前依旧有一些无法解决的问题，其中之一便是拦在从自分裂的脂质分子演变到真正的原始细胞之间的首要问题：如果细胞内的RNA的复制快于细胞生长，那么细胞会长到足够大再进行分裂，但如果是细胞生长快于RNA复制，那么RNA会渐渐变得不足，新生细胞中将出现没有RNA的空壳囊泡。为了能够生存，生命必须平衡两者，精确调节复制和生长之间的关系，以便使RNA的复制不快于细胞本身的生长。这种协调性到底是如何建立的，是20世纪科学遗留给后人的问题之一。



让我们从牛车直接快进到法拉利。虽然生命的某些特征在它们出现之后的3 000多万个世纪里都没有改变过，我们将在后续的章节里看到，生命的成分分子、调节方式以及新陈代谢一直都是新性状出现的源泉，但是进化也在不断塑造着生命除此以外的方方面面。早期原始的RNA复制体变成了复杂的蛋白质酶系，除了RNA和脂质，生命还学会了调节和平衡数千种其他分子。无数后来出现的生化反应将现代细胞的新陈代谢，相当于法拉利的引擎，变成了一项化学技术上的奇迹。

想象一下，你开着这辆法拉利从一场晚宴上回家。时值深夜，却在高速公路的某处发现燃料耗尽，目之所及没有任何加油站，也没有顺风车可以搭。但是没有关系，你打开后备厢，里面的冰箱里还有一些剩余的食物和饮料。你向油箱里倒了一瓶橙汁、一升牛奶和一杯酒。这些足

够让你渡过难关，把你送到下一个加油站了。于是你又重新上路。

现代的新陈代谢过程正如上述的法拉利引擎，它们能够利用许多不同种类的燃料。除了燃烧供能之外，新陈代谢还可以利用所有这些燃料获得并合成身体所需的基本粒子，身体会利用这些粒子进行生长、繁殖或是修复伤口。这就好比一辆车不光能够利用油箱里的燃料启动引擎，同时还能用它修补漏气的轮胎和破损的挡风玻璃。

我们这里所说的基本粒子包含大约60多种核心分子，它们是构成以及修复人体的主要成分。最重要的基本分子莫过于组成DNA的4种脱氧核苷酸，也就是构成人类基因组的单位成分。每个脱氧核苷酸分子由一分子脱氧核糖、一分子磷酸基团以及一个含氮碱基构成。含氮碱基一共有4种，分别为腺嘌呤、胞嘧啶、鸟嘌呤及胸腺嘧啶。紧随其后的是DNA的转录产物RNA，同样是调节生命活动的重要分子。组成RNA的4种核苷酸分别为腺嘌呤、胞嘧啶、鸟嘌呤和尿嘧啶（uracil，U），和组成DNA的脱氧核苷酸仅有一个氧原子的区别，不过正是这个氧原子导致了巨大的化学差异，使得RNA更适合作为催化剂。

由于缺乏氧原子的核糖更稳定，所以DNA更适合作为遗传信息的载体。RNA继而被翻译为蛋白质，构成蛋白链的基本单位是20种氨基酸，其中的一些在日常生活中十分常见，比如感恩节后嗜睡症的元凶色氨酸，还有调味剂味精的主要成分谷氨酸。除此之外，还有生物膜的主要成分磷脂，在食物不足时的能量储存分子，协助酶完成催化作用的分子等，正是类似的大约60种单位分子构成了细胞本身。

新陈代谢的主要任务在生命出现的38亿年间几乎丝毫未变，主要是获取能量以及合成物质。新陈代谢反应本身也没有改变，以前一分子蔗糖通过水解反应得到一分子的葡萄糖和一分子的果糖，现在依旧如此，改变的仅仅是新陈代谢反应的数量。我们远古的祖先只需要依靠寥寥几个生化反应就可以活命，而现代生物则要依赖众多复杂的新陈代谢反应。

现代的新陈代谢是一系列高度复杂且相互关联的生化反应组成的网络，这张反应网是生命经历将近40亿年进化的结果。如果你试着把这些反应绘制出来，它看起来像极了一张标注出每条街道的美国地图。从居民区的小巷到整条州际高速，一切尽收眼底。图的中心是古老的三羧酸循环，就像连接白宫和国会大厦的宾夕法尼亚大道。图2-2展示了这张反应网络的一小部分，图中以线条相连的两种物质（在图中以图形表示）之间都存在相互反应。你可以把它当作一张村庄的地图来看，图中标出

了蔗糖分解反应中的4种相关分子，它们都被圈在一个椭圆内。不过，不要被这幅简化图欺骗了，它所展示的并不是完整的事实。果糖实际上在人体内参与了37种不同的反应，而不只是图中展示的这一种。另外，现代的新陈代谢反应需要底物以外的许多其他分子参与才能完成。

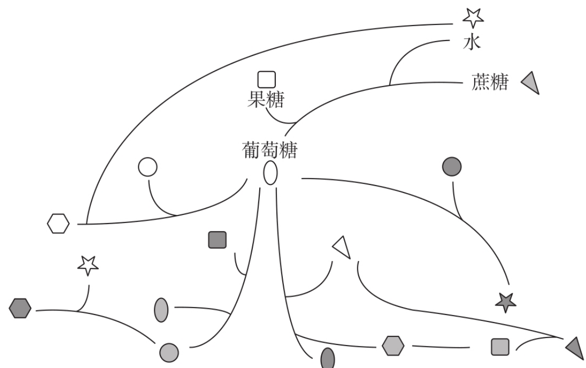


图2-2　部分新陈代谢网络示意图

弄清代谢反应网络中涉及的分子花费了科学家一个多世纪的时间。在过去的100多年里，数以千计的生物学家通过研究同一种人类肠道细菌构建了有关新陈代谢的知识巨塔，这种细菌就是大肠杆菌。科学家构筑这座知识巨塔耗费的时间和精力几乎与在现实中建造一座中世纪大教堂无异，而从塔顶看到的风景也蔚为壮观。

如今，我们已经意识到大肠杆菌的新陈代谢十分奇异，包含数百个代谢反应以及反应中涉及的数千种分子。我们也意识到，就新陈代谢这方面而言，大肠杆菌以及许多其他微生物都可以轻易打败我们。比如，对于组成蛋白质的20种氨基酸，我们的身体只能合成其中的12种，其余的氨基酸只能通过食物获得；正常情况下身体需要的13种维生素，只有两种是我们的身体能够合成的，即维生素D和B_7（生物素）。而大肠杆菌可以从零开始合成所有这些维生素。

大肠杆菌的新陈代谢之所以如此复杂，关键在于我们上述所说的那60种生物单位分子。合成每一种基本分子都需要众多相关反应以及中间产物，而大肠杆菌是一名出色的生存游戏玩家，营养丰富的肠道并不是它最得意的竞技场，哪怕是贫瘠到只有7种小分子的饥荒环境也可以是它们的乐土，它们依旧能够利用这些分子获取能量和营养。在这种极端的环境里，每分子物质都必须身兼两职，比如葡萄糖就在为大肠杆菌提供能量的同时也为它的合成代谢提供碳元素。大肠杆菌仅凭这些就可以

合成任何需要的基本物质，然后再用这些单位物质获得其他所有所需的生物成分。

作为一名生存型选手，大肠杆菌的本事还不止于此。如果从已经贫瘠不堪的环境里取走所有的葡萄糖并替换成另一种不同的物质，比如甘油，大肠杆菌依然能够利用这种新的成分为自己提供碳元素和能量。把甘油换成醋酸，道理也相同。总共算起来，大肠杆菌可以利用超过80种不同的分子作为它唯一的能量以及碳原子来源，进而合成细胞内的千万亿分子。对于其他几种元素也类似，比如氮元素和磷元素。大肠杆菌就像一台能够自我构建、自我增殖、自我修复的跑车，而它需要的燃料既可以是煤油，也可以是可口可乐，甚至可以是洗甲水。

成分越是简单的化学环境越适合微生物的实验室研究，但在自然界中如此纯粹可控的环境往往不常见。在类似土壤和人体肠道这样的环境里，物料分子的种类总是不断发生着变化。为了从这样的环境中有效摄取能量和碳源，微生物代谢的物质需要有一个明确的先后顺位。而要建立这种顺位，它们就必须尝试每一种可能的能源和碳源。

这样一想，1 000多种反应听起来似乎也不算多了。

当今的生物与它们遥远的祖先的另一个重要区别在催化剂，也就是加速化学反应的功能分子。如果你的肠道缺乏适当的酶，比如蔗糖酶，那么你可能要花上几年甚至数十年时间才能消化一杯糖水里的蔗糖。如果没有蔗糖酶的帮助，就算你每天喝十几升的糖水，最后依旧会死于低血糖。

不过，现代生物的催化剂已经不是简单的金属元素催化剂了。如今自然界的生物催化剂可以成万亿倍地提高生化反应的速度，让底物分子几乎在相遇的同时就完成反应。自然界有数千种不同的催化分子，每一种都有特定的氨基酸序列。再以蔗糖酶为例，蔗糖是一个包含1 827个氨基酸残基的巨大分子，每一个氨基酸残基至少有十几个原子，也就是说一个蔗糖酶分子里有两万多个原子，但是蔗糖分子总共只有45个原子。与蔗糖酶相比，如果说蔗糖是一粒豌豆，那么蔗糖酶就相当于一个足球，这也就是为什么相对于它们所催化的底物或者合成的产物而言，酶分子会被称为生物“大分子”（macromolecules）。蔗糖酶看起来已经不算小了，但是类似大小的酶在自然界比比皆是，很多酶的尺寸甚至远远超出于此。

蔗糖酶的氨基酸链合成之后，需要进行空间上的弯曲和折叠，如同

毛线球一样，但是两者有一个重要的区别：每个毛线球可能都略有不同，但是每一个蔗糖酶都完全一样。蔗糖酶的氨基酸链合成之后，会在空间上以严格的方式进行精确的折叠。经过折叠的蔗糖酶通过高频的扭曲、摇摆和震动执行它的催化作用。我们可以想象一下这台自我组装的纳米机器，它行动迅速地吸收底物分子，裂解之后吐出反应产物，整个过程一气呵成，快得让人眼花缭乱。

每一个细胞都含有数千种类似的纳米机器，每一种都负责催化一个特定的生化反应。所有这些酶都在细胞内生物单位分子高度集中的区域内发挥作用，这些代谢反应发生的特定位置通常比东京高峰时段的地铁站还要拥挤，令人称奇。



我们还不知道生命到底是如何从最初简单的形式进化出如此高度的复杂性，或许我们永远也无法知道确切答案。到目前为止，在化石中发现的最古老的细胞已经与现代细胞无异，而它们的祖先至今仍然半遮着容颜，隐藏在氤氲之中。这种未知一点都不奇怪。多数古老的岩石都无法在漫长的时间长河里保留下来。最早的原始生命不过是一团柔软脆弱的分子，即使动荡的大陆板块没有把它们留在岩石上的痕迹抹得一干二净，它们也不是铺满海底的蓝绿藻（blue-green algae）[10]，更不用说像生活在数亿年前的恐龙那样，留下巨大的骨骼化石。

但我们可以确信的是，所有生物都来自一个共同的祖先，这并不是说生命起源只发生过一次。由于自组织现象的存在，我不会对历史上生命有过多次起源感到惊奇，最早的生命可能诞生于深海热泉，可能诞生在温暖的池塘，又或者，天晓得是哪里。在所有这些忽明忽暗闪烁于地球早期的微弱的生命之光中，有的星火难以为继，有的则越来越明亮。它们之中只有一个得以辉煌灿烂，并诞下了今天所有的生命。这不是“仁者见仁，智者见智”的问题，而是必须如此，原因只有一个：标准化，精确并且广泛适用的标准化。

计算机学家安德鲁·塔嫩鲍姆（Andrew Tanenbaum）曾经不无嘲讽地说：“标准化的唯一好处是，它让你有充足的选择余地。”我大概明白他所嘲讽的对象。每当我家里的遥控器、钟表或者别的什么小玩意儿没电的时候，我就要翻箱倒柜地找出一大把大大小小的电池，但通常都没有我需要的型号。如果日常生活中只存在一种规格的电池，抑或只有一

种型号的咖啡滤纸、数据存储介质和操作系统，那不知道要免去多少麻烦。甚至更古老的技术都头疼于难以统一的标准：在公共电力系统建立一个多世纪之后的今天，世界上依然存在14种互不兼容的插座标准。每天，当全世界上百万个国际旅行者带着笔记本电脑、电吹风和剃须刀到达一个陌生的城市，却发现忘记带上合适的插座转换器时，想必都是万般无奈。

大自然不一样，它有标准化的电池，有着各种可利用的能量形式，包括机械能（拆迁时用铁球撞毁房屋）、电能（驱动电脑的电子流）和化学能（分子中把原子连接在一起的键能），其中化学能是最受生命青睐的。地球上的所有生物，从单细胞的细菌到巨大的蓝鲸，都使用同一种标准化的储能物质，这种能量分子就是三磷酸腺苷（adenosine triphosphate, ATP）。三磷酸腺苷分子中有高能的化学键，当高能化学键断裂时，键能就会转移到其他分子中，同时三磷酸腺苷变为相对低能的二磷酸腺苷（adenosine diphosphate，ADP）。为了重新合成三磷酸腺苷分子，需要某些特殊的酶催化，将能量从能源分子转移到二磷酸腺苷当中。

不过，并不是所有来自三磷酸腺苷的能量都会被转移到其他分子上。细菌用三磷酸腺苷的能量挥动鞭毛，驱动自身在水里游动。萤火虫则在希望吸引配偶的时候用三磷酸腺苷点亮自己的身体。有些种类的鳗鱼会把三磷酸腺苷的化学能转化为电能，并用电脉冲捕捉猎物。但是无论最终变成什么形式的能量，不管是机械能、光能还是电能，生物利用的所有能量本质上都是来自三磷酸腺苷的化学能。

如果细胞想利用能源物质来合成自身的成分，比如葡萄糖，它必须首先将葡萄糖里的化学能转移到三磷酸腺苷里。而后一步接一步，三磷酸腺苷的化学能被用于合成其他分子。通过这种方式，来自食物的化学能最终成为生物成分分子中的化学键能。因此，三磷酸腺苷是能量转移过程中关键的中间分子。

所有生物都以三磷酸腺苷为通用的标准能源物质，它们不需要检查电池的型号，也不用在机场为插座转换器支付额外的溢价。现存的所有生物都继承了某个祖先发明的储能标准。然而，这个出色的标准化能源并不是生物唯一的标准化项目。我们已经见识过新陈代谢的中心反应三羧酸循环了，还有自然界通用的生物膜里的脂质分子与水的爱恨情仇。除此之外，还有DNA、RNA以及每三个核苷酸分子对应一种氨基酸的密码子编码方式，所有生物都采用同一套密码子。

三磷酸腺苷和三羧酸循环作为生物界的通用标准，与光速作为宇宙速度的极限存在些微差异。三磷酸腺苷和三羧酸循环不是生命唯一的选择。我们已经发现了可以遗传编码的潜在方式，还有能量载体三磷酸腺苷，甚至是作为遗传信息载体DNA的可能替代物。所以，生物体的标准化不是必然，而是某个远古的共同祖先的遗留物。生命起源之初，有许多踌躇满志的选手对这场进化的马拉松跃跃欲试，不过由于自然选择或者运气不佳，最终只有一名选手坚持到了终点线，留下了自己的子嗣。如果设身处地地体会一下这个祖先过关斩将、披荆斩棘最终子孙满天下的过程，个中艰辛不禁令人感到些许绝望。所幸，“祸兮福之所倚”，从中我们得到的启示是，至少对于常年旅行的人来说，或许再等上40亿年，人们就不再需要插座转换器这种烦人的东西了。



当你读到这里的时候，你已经对生命起源的谜题有所了解了。我们现在知道，生命可能起源于温暖的“小池子”，可能起源于深海热泉，也可能起源于冰封的大海，甚至起源于外太空。也许我们要再等上一个世纪才能知道答案。不过，对于理解生命起源和进化来说，相比于弄清实际的过程，有两个启示在现阶段显得更为重要。

第一个启示，生命需要进化的能力，甚至在生命还没出现的时候就需要，以保证自催化的新陈代谢和最早的自我复制体诞生。

第二个启示，生物进化的交响曲有三段不同的主旋律。第一个篇章，进化把不同的化学反应组合到一起，比如合成生物单位分子的代谢反应以及合成第一个自我复制分子。第二个篇章，进化需要借助促进分子反应的辅助分子的力量。第三个篇章，进化创造了调节，这是高度复杂的生命体维持自身稳定的关键。伴随着生态圈的生命体变得越来越复杂，适应力不断增加，进化的这三个主旋律回荡在历史长河里，振聋发聩。

原始的新陈代谢演变为复杂的反应网络，网络中的反应不断发生重新组合，让生命尽可能地拓展到了任何可能的栖息地中。复杂的蛋白质酶替代了无机催化剂，并让功能复杂的蛋白质的出现成为可能，比如感光用的视蛋白以及防御用的角蛋白。还有调节，虽然它看起来似乎无关紧要，却是进化必不可少的组分，正是由于调节过程的存在才让多细胞器官得以出现，如四肢、心脏和大脑。

从生命出现到今天，进化一直在不断改变和优化新陈代谢、蛋白质和调节。虽然这三者看起来毫无联系，但在它们背后起关键作用的，正是神奇而强大的自组织形式。

ARRIVAL OF THE FITTEST

03 宇宙图书馆

Solving Evolution's Greatest Puzzle

一种生物所具有的全部生化反应构成了这种生物的新陈代谢。新陈代谢进化的本质在于重新组合。生命时刻在尝试每一种可能的基因新组合，重新解读，重新编译，然后重新布局代谢遗传，毫不停歇，从而造就并提升着代谢的多样性。新的代谢能力是不断驱动生命拓展最前沿阵地的引擎。



想象一下，你站在一个堆满书的房间里，书垛直冲天花板。四面的墙壁上都是成排的书架，连留个门的位置都显得够呛。你穿过房间，开始翻阅周围的书。很快，你就发现房间里每一本书的页数，每一页中的行数，以及每一行里的字数都不多不少，全部相同。不过奇怪的是，这些书中的内容犹如痴人的呓语，不知所云。每本书的每一页，每一页中的每一行都是字母的随机排列，诸如“hsjaksjs……”或者“zvaldsoeg……”等，凌乱无序的字母中偶尔穿插着空格和标点。只有在十分难得的情况下，你才会找到几个有意义的单词，比如“cat”（猫）、“teapot”（茶壶）、“bicycle”（自行车），它们就像漂浮在文字垃圾海洋上的鲁滨孙之岛。

不消多时，你肯定就会对这些毫无意义的书感到厌烦。于是你选了其中一面墙上的门想出去透透气，推开门却发现自己进入了另一个一模一样的房间：四面墙上各有一道门，每扇门旁都围着密密麻麻的书架。书架上的陈列依旧如同天书，毫无意义可言。

这个房间里的门又把仍不死心的你带到了另一个几乎一模一样的房间，一个接着一个，无穷无尽，直到你终于意识到自己身陷于一个没有尽头的迷宫里，除了成堆的书，周围的一切都一模一样。你在探索的途中遇到了其他人，从他们嘴里你得知这个藏书的地方巨大无比。难以计数的书构成了这个庞大而又神秘怪异的图书馆。

我们姑且把你身处的这个房间称为“宇宙图书馆”，里面收纳了世间所有的书籍。

确切地说，所谓“所有的书籍”是指所有字符的所有组合方式，即26个英文字母以及标点符号的所有组合。这种随机组合方式的典型产物你已经见识过了，正是图书馆里那些毫无意义的文字垃圾。不过，偶尔你也可以在某本书里找到一个有意义的单词，一个表意通顺的句子，甚至是一整段话。按照这个思路，可以想见在图书馆的某些角落里，我们还是能够找到一些符合语法、言之有物的书。由于宇宙图书馆里收录了所有可能的书，这也就意味着它收录了所有在人类历史上被撰写和出版的书。

所有可能被书写的小说、短篇故事、诗集、真实或虚拟的传记、哲学专著、宗教典籍、科学及数理论著；除了用英语撰写的书，甚至还有用任何文字书写的书；有阐释世间真理的书，也有散布虚伪谎言的书；

有对于其他书进行评论的书，有关于这座图书馆前世今生的书，有的正确，有的谬误；有关于你一生的书，告诉你你的人生从何开始，又将去向何方、归于何处；当然，也包括你现在正在阅读的这本书。所有这些书都被收录在这个图书馆中，宇宙图书馆的规模远远超出你的想象。

如果我们想对宇宙图书馆的规模大小形成一个模糊的印象，不妨假设馆里的每本书里有50万个字母（这不算特别多，基本和你正在读的这本书相仿）。不考虑标点符号，50万个字符中的每一个仅有26种不同的字母选择（从A到Z）。具体说来，一本书的第一个字母有26种可能，第二个字母依旧有26种可能，而后第三个、第四个……如果要计算有多少种可能的书，我只需要计算26的50万次方，也就是说$26^{500\ 000}$。这是个非常巨大的数字，在1的后面足足跟了70万个0，光是这些0的数目就已经比书里的字母多了。这个数字是一个超宇宙常数，已经远远超过了宇宙中氢原子的数量。

宇宙图书馆里的馆藏就是自然母亲创造力的真相：全包全揽、无穷无尽。只不过，我们在宇宙图书馆中要讨论的并不是用人类的文字写就的书籍，而是用遗传字母和化学分子谱写的DNA。



人类的文字或许能够记录整个宇宙，前提是那些语言可以涵盖的内容，但在这座宇宙最古老的图书馆里，化学才是创造新陈代谢和生命的通用语种。人类可以用散文和诗歌歌颂这个星球上数以万亿计的任何生命，但创造这些生命却只能用化学语言，特定的化学反应遇到生命基本的构成物，继而造就生命体。图书馆里的所有化学语言之和就是生命之歌。

在第2章中曾经提过，我们在地球上的部分生物体内已经总计发现了超过5 000种合成生物体自身物质的化学反应，包括用来合成DNA和RNA的核苷酸，以及用以合成蛋白质的氨基酸。大肠杆菌细胞内发生的近1 000个生化反应，正属于这个范畴。此外，还包括所有细菌、真菌、植物、动物及人类体内的化学反应。多亏了这些化学反应的存在，人类的身体才能够从糖和其他食物中吸收能量，修复不小心摔破的膝盖，补充身体里每天损耗的数百万个红细胞。

没有哪种生物可以同时具有所有的5 000多种生化反应，每一种生物只能利用其中的一些，一种生物所具有的所有生化反应就构成了该生

物的新陈代谢。多亏了20世纪生物化学领域的新发现和21世纪早期的技术革命，我们才能通过对众多物种的研究，从而了解这些反应。目前，科学家已经把超过2 000种生物的代谢信息储存在巨大的在线数据库中，如京都基因与基因组百科全书（Kyoto Encyclopedia of Genes and Genomes）以及BioCyc数据库。任何接入互联网的计算机都能方便快速地访问这些数据库。

图3-1代表了一种我们如何组织这些信息的方式。左侧列出了5 000种不同的生化反应，每个生化反应都以化学方程式的形式表示。为了避免冗杂，我只写出了其中的一个方程式：蔗糖的分解反应。其余的反应物都以简单的字母替代。我们考虑某种特定的生物，比如大肠杆菌或人类，如果这种生物体内具有该反应，我们就在对应的方程式右侧标记一个“1”，代表它具有相应的基因，负责编码催化该反应所需的酶。否则，我们就标记一个“0”。于是便得到了一长串连续的“1”和“0”，正如图中所示的那样，我们可以用这串由“1”和“0”构成的数列代表任何一种生物的新陈代谢模式。

	生化反应			基因型
5 000种生化反应	蔗糖+水	→	葡萄糖+果糖	1
	A+B	→	C+D	0
	⋮			⋮
	D+E	→	F+G	0
	H+I	→	J+K	1
	⋮			⋮
	I+K	→	L+M	1
	N+P	→	O+Q	0

图3-1　代谢基因型示意图

像大肠杆菌这样的细菌可以合成所有20种构成蛋白质的氨基酸，而像人类这样的代谢“差生”则只能合成其中的12种。我们缺乏合成其余8种氨基酸的酶和化学反应。以图3-1中的简化法描述新陈代谢可以很形象地解释物种间的区别：由于我们缺乏相应的生化反应，对某些生化反应而言，我们的标记就是“0”，而大肠杆菌的标记是“1”。

这种数列相当于一种简化的生物代谢基因型，是所有代谢反应的总和，也就是新陈代谢，所以代谢基因型包含了一种生物基因组中与代谢

有关的所有基因。你可以把它看作是一种用二进制书写的文本，既没有标点，也没有空格，譬如“1001…0110…0010”。文本的第一个字符代表蔗糖分解反应，这里它的编号是“1”，而第二个反应可能代表合成某种必需氨基酸的反应，在这里的编号为“0”，代表这种生物不具有这种反应，而另一种生物则可能具有，也就是说其基因型编号是“1”，以此类推。

上述文本只是宇宙图书馆里的其中一个例子，事实上，庞大的图书馆内包含了所有可能的代谢基因型。

用计算图书馆书本数量的算法，我们同样可以计算这种编码方式下的所有编号数量。每一种生化反应对于某种特定生物的新陈代谢来说只有两种可能性，存在或不存在。对于第一个反应有两种选择，第二个反应亦然，以此类推。当检验过每一种生化反应后，编码的总数就等于与反应数量相同个数的2相乘。就已知的5 000个生化反应而言，可能的基因型一共有$2^{5\ 000}$种，每一种基因型都是由“0”和“1”构成的数列，代表一种不同的代谢种类。这个数字超过$10^{1\ 500}$，也就是1后面跟着1 500个0。虽然比不上我们上文中的书本多，但也已经远远多于宇宙中的氢原子数了。代谢图书馆内的馆藏数量同样超乎常识。

如同随机庞杂的宇宙图书馆里包含了所有真实存在的书，代谢图书馆里同样包含了所有“真正”的代谢基因型，即那些真实存在于某种生物体内的代谢模式，而另一些并没有实际意义，只不过是乱码的书本而已。有的代谢基因型无法令生物获得能量，而有的则无法合成重要的代谢物质。好比一本书，虽然有的章节、段落或句子语意通顺、语法正确，但整本书却没有主旨，逻辑混乱。更有甚者，通篇连一句有意义的句子都难得一见，只有混乱无序的字母串。这些基因型所代表的代谢由缺乏关联的生化反应组成，它们的合成反应往往以对生物无用的产物分子大量囤积而告终。



如果你在宇宙图书馆里停留足够久，一定会发现一些在主旨、想法和创意上让你颇感惊讶的书。代谢图书馆里的馆藏在这方面也是一样。你会发现前所未有的生化反应、合成新颖小分子的表现型以及利用新能源的能力。换句话说，你会发现一些新的性状。

新陈代谢与生物进化几乎一样古老，不断进化的生命几乎一经诞生

就开始探索这座庞大的图书馆。大自然早在10亿年前就创造了数量多得难以想象的生物性状，远远超出了实际需要。然而进化并没有因为这些早期的成就而骄傲自满、停滞不前。在数以万亿计的现存生物中，新的生物性状依旧以远远超出我们解读能力的速度不断涌现。某些新性状出现的时间还不到100年，对于整个进化史来说，这仅仅是一瞬间而已。

我们来认识一下五氯苯酚，人类第一次学会合成这种臭名昭著的分子是在20世纪30年代。它被作为防污涂料用于船体表面，同时也被作为杀虫剂、除真菌剂以及消毒剂。简而言之，五氯苯酚被用来杀死各种生物。五氯苯酚对人的肾脏、血液以及神经系统同样有害，此外，它还是一种致癌物质。不过，即使它剧毒无比，生命还是找到了方法耐受它的毒性，甚至把它作为美味佳肴。鞘脂菌属的细菌S.chloroplenolicum[11]，顾名思义，能够利用五氯苯酚同时作为自己的碳源和能源，并且五氯苯酚是它唯一的食物来源。为此，它的基因组编码了4种催化用的酶，用以将五氯苯酚转化为像葡萄糖一样容易消化的分子，这相当于把生化武器变成了自己的战争口粮。

这种利用五氯苯酚的能力是S. chlorophenlicum特有的，但代谢的化学反应本身却不是。五氯苯酚代谢过程的每一步反应都可以在其他数百种乃至数千种生物体内找到。其中两步反应在某些细菌中起到循环利用多余氨基酸的作用，而其余的两步反应则会参与分解某些真菌和昆虫分泌的毒性分子，因为这些毒性分子的结构恰巧和五氯苯酚类似。进化就像一座由自动报警的洒水系统、气泵和聚氯乙烯管等组合而成的机械停车楼，它利用不同生物中已然单独存在的各种反应，重新组合出了独特的S. chlorophenolicum。也就是说，新陈代谢进化的本质在于重新组合。

生物体通过进化获得摄食人造剧毒分子的能力，这种现象在自然界并不鲜见。伯克氏菌属的细菌B.xenovorans能够大啖多氯联苯，而这种曾经被广泛应用在塑料制造和电气工业领域的化合物已经被明令禁止。还有一些细菌甚至能消化氯苯，后者是化学实验室普遍使用的一种剧毒有机溶剂。更极端的是，有的细菌甚至可以分解和吸收专门用来杀死它们的抗生素。能被细菌作为食物的抗生素中包括一些人造的种类，所以它们利用这些抗生素的历史并不长。

自然力量不仅能为无米之炊，把毒药变成生命的美味口粮，还能贤惠地废物利用。以氨气（NH_3）为例，你可能觉得它不过是家用清洁剂里刺鼻而难闻的那种气体，但它除了辣眼睛之外，还是一种剧毒的动物

代谢产物。由于氨气易溶于水，所以鱼类可以直接把代谢的氨排入周围的水里，而后扬长而去。对于人来说，这就好比是排尿的过程。然而当3亿年前动物开启进军陆地的征程时，它们再也享受不到这种随时如厕的福利了。陆生动物亟需一种新的方式排出血液中的剧毒氨气。

这种新的方式可以在代谢图书馆里找到，那就是把氨气转化为毒性较低的尿素，直到今天，尿素依旧是我们尿液里的主要成分。尿素的合成反应包含了五步普通化学反应，远在削减氨气毒性的反应之前。尿素合成反应中的每一步反应都已经在不同生物体中存在，互不相干，井水不犯河水。

我们不知道动物学会合成尿素的确切时间点，不过相关的线索俯拾皆是。虽然现代多骨鱼，即硬骨鱼，不需要用转化代谢的方式来降低氨气的毒性，但是作为硬骨鱼的祖先，同样游弋在海洋里的软骨鱼早在硬骨鱼出现之前就已经学会合成尿素了，代表鱼类有鲨鱼和鳐鱼。大白鲨合成尿素的目的与人类稍有不同：它们不仅利用尿素作为氮元素的储备池，同时还用尿素保持自身的浮力和在海水中的平衡。你可能会想，如果硬骨鱼遥远的祖先能够合成尿素，那么在它的DNA中是不是可以寻得一些与合成尿素有关的蛛丝马迹。倘若如此，你的确没有想错：主导尿素循环反应的基因的确还存在于硬骨鱼当中，只不过它们在绝大多数情况下都不表达。这些沉默的基因在硬骨鱼体内就像牙牙学语时的我们，虽然能够认得些许词汇，却也是有口难言。

清除垃圾不如废物利用，而大自然尤其擅长后者。无论是氨气还是尿素，动物排出的含氮废物都是植物的肥料。而我们呼吸的每一口氧气也不过是植物光合作用产生的“废物”。每一克动物排泄物里都含有数十亿个细菌：人类排出的废物恰恰是这些微生物的无价之宝。粪便里的每种细菌都有自己独特的代谢方式，不管代谢模式是新是旧，都可以用于降解粪便里的有机分子，为细菌提供能量和所需的分子，使它们繁荣昌盛、生生不息。

代谢的进化不仅发生在适宜的环境里，在极端环境中也同样常见，如极端高温、极端寒冷、极端干燥、高度腐蚀性、辐射过量、极度高渗等。细菌作为个中典型，能够在沸腾的水里生息，也能在冰天雪地里泰然自若，既不害怕具有腐蚀性的硫酸，也对有着致命压强的深海毫无畏惧。为了能够在这些环境里生存下去，它们经历了无数次进化，而许多进化都与代谢相关。

如果没有这些进化，极端环境可以像这些细菌杀死我们一样，轻易

地让细菌们毙命。以高盐环境为例，由于酶在执行自身功能时依赖水作为溶剂，高盐环境中的高渗透压可以令细胞脱水而死。为了弥补损失的水分，代谢进化出了一些独特的物质，比如四氢嘧啶和甜菜碱。这些名字古怪的分子没有水那么容易脱离细胞，能够在水分顺着渗透势离开细胞的时候作为水分子的替代物。它们可以维持蛋白质的溶解状态。而合成这些分子仅仅需要几步额外的化学反应，以及一些常见的物质作为原料，比如天冬门氨酸盐。把这些合成反应整合到你体内的新陈代谢中，你就获得了在相应的极端环境中立足的资本。嗜盐菌（halophilic bacteria）——它的名字来源于希腊语“喜盐”（salt-loving），能够在浓度高达30%的高盐环境里存活，10倍于人类所能耐受的极限浓度。嗜盐菌能够在盐晶体周围甚至晶体内部存活。

与其他生物比起来，可怕的极端环境倒显得有些不值一提了。掠食者和捕食者都是生物生存的大麻烦，尤其当你无从逃避的时候。由于无法移动，常见的植物基本都是其他生物的刀下肉，如昆虫、生活在地底的蠕虫、地面上的蛞蝓和食草动物都把植物当作盘中餐。植物无法通过行动进行防御，所以它们进化出剧毒的化学物质令动物避之不及。植物并不是这场化学战争里的唯一参与者，但确实是个中精英和翘楚，其中的原因大概正是因为它们哪儿也去不了。

毒性分子的合成需要植物整合特定的生化反应，所以这些防御性分子均来自植物经历的长期进化。其中一种分子名叫尼古丁，也是令许多人吸烟时如痴如醉的烟草植物合成物。由于其巨大的毒性，尼古丁也被一些农民用作杀虫剂。但最近一组德国科学家发现，植物才是这种杀虫手段的首创者。当他们人为地降低烟草植物内的尼古丁含量后，某些害虫开始对它们大快朵颐。这些昆虫对烟草的攻击更频繁，吞噬的叶片更多，生长更迅速。而对于烟草而言，它们在掠食者的攻击下失去了比普通烟草多3倍的叶子。

尼古丁只是我们现在已知的3 000多种植物碱里名声最响亮的那个。植物碱指一大类围绕氮原子构建的有机分子，包括咖啡因和吗啡，它们是植物的化学自卫武器。此外，虽然种类繁多，但植物碱也只是植物众多化学武器中的一种。其他的“化武”还包括涩味的丹宁，也就是食用不熟的水果时让你的嘴巴感到干涩的罪魁祸首。丹宁会与植物的蛋白质紧密结合，阻止它们在我们的肠道内被消化，这使动物对合成丹宁的植物心生厌恶而不愿意优先摄食它们。

最为臭名昭著的是一种叫生氰糖苷的化学防御物质，主要存在于非

洲和美洲的主要粮食作物木薯和树薯中。如果不经由充分烹煮与浸泡除掉生氰糖苷，这些作物就会释放氰化氢，也就是齐克隆B[12]中的活性成分，后者曾经被泵入纳粹奥斯维辛集中营的“洗浴室”里。如果你还在幻想大自然是一个诗情画意的秀丽之地，是伊甸园里的后花园，那么植物的生化武器可以立马把你天真可爱的愿景轰得灰飞烟灭。

上述生化武器分子都是对已有化学反应重新组合得到的产物，新的反应顺序让普普通通的原料转化为剧毒物质。反应的每一步都需要代谢基因型中一段特定的文本作为指导。



不同物种获得新代谢的方式十分类似，这些方式在大型的多细胞生物中也很常见，人类就位列其中。表现之一就是伴随着有性生殖出现的性状改变，有性生殖后代性状变化的原因主要是来自亲本的染色体发生随机组合和重新洗牌，所以我们每个人都是从异于父母的起点开始各自的生命旅程的。此外，DNA还会由于一些随机事件发生自发突变，包括紫外光子冲击以及代谢过程中产生的高能氧自由基损伤DNA分子中的化学链接。

由于有性生殖的“重新洗牌”只发生在高度相近的基因组之间，而任意两个人的基因组相似度都高达99.9%，所以上述两种检索方式在代谢图书馆里都算不上高效。打个比方，如果你只修改《哈姆雷特》中的30个单词，并不能把它改成一部全新的作品。另外，虽然变异可以创造出新蛋白，包括新的催化酶，但这种概率非常小，意味着纯粹依靠变异的进化过程将十分缓慢。

此外，代谢进化在大型、多细胞动物中进展缓慢还有一个原因。有价值的能量获取新方式和生物体新结构在种群中的传播范围与传播速度正相关。对于生殖周期为数十年，哪怕是数个月的动物来说，由于繁殖速度的限制，它们的种群都无法快速地实现进化。

即使面对无数的不利条件，包括人类在内的动物在代谢进化方面也并不是无所作为的。我们的身体能够降解药物，比如生活中常用的阿司匹林，化学家则称之为乙酰水杨酸。通过一种叫葡萄糖醛酸结合反应（glucuronidation），阿司匹林可以被修饰为毒性较低的产物继而随尿液排出。猫、鬣狗等掠食动物体内则缺乏这种代谢需要的酶。（所以下次在给你的宠物狗喂阿司匹林之前，最好先咨询一下你的兽医。）你可

能会问，远在拜耳公司把阿司匹林这种药投入市场的20世纪80年代之前，我们的身体为什么要在进化中保留这种酶呢？回答这个问题的线索在阿司匹林的名字本身，“aspirin”取自一种绣线菊属植物——榆绣线菊（spiraea ulmaria）。这种植物和许多其他植物在很早以前就被用于止疼。不仅如此，含有水杨酸的植物曾是我们祖先采集的食物之一，因此，与鬣狗那样纯粹的食肉动物不同，作为杂食动物的我们需要一种降解水杨酸毒性的手段。

不过，在多细胞生物的世界里，人类根本算不上代谢竞技擂台上的种子选手，许多动物在代谢的不同方面都胜于我们。人类无法合成维生素C，所以许多人早餐时都要来一杯橙汁，而狗却能够合成自身所需的维生素C。虽然我们能从植物的种子，如大麦和玉米中吸收热量，而奶牛则可以消化和吸收植物茎秆中的纤维素。说句公道话，追根溯源，消化纤维素的神奇能力并不是奶牛自己的本事，而是由于它们体内的微生物：牛的4个胃里的细菌能够将巨大的纤维素分子分解成易于消化的葡萄糖。

这似乎暗示我们，进化的真正好手其实是我们星球上最小的生物：细菌。

细菌拥有强大的繁殖能力，它们的生殖周期只有数分钟，因而基因库的更新速度远快于我们。但是细菌具有的进化优势远远不止于此。为了让你能够理解人类和它们的巨大差距，我们可以想象有一个身高只有1.5米的小伙子，他一直希望能够加入高中的篮球校队。努力的锻炼和勤奋的练习对他的帮助杯水车薪。他的最大问题是没有合适的基因，而他最好的朋友只要踮起脚就几乎能够碰到篮筐。

而对于细菌来说，如果一个细菌想要与另一个细菌比肩，它们的出身可不是决定性因素。如果我们在这里讲的是一个科幻故事，这对好朋友拥有了和细菌一样的进化能力，那么你接下来将看到的一幕是：当这两个小伙子在他们喜爱的一家饭店吃饭时，一根细长的空心管子从高个子的体内伸出，摸索着伸向矮个子小伙儿。一旦两人被连接在一起，这跟管子随即把高个子的一小块DNA片段注入到矮个子体内。如果注入的片段中正巧包含了与身高有关的基因，那么学校的篮球队就有了一个新的大前锋。

这是一个基因水平转移的例子。可惜的是，落后的人类并不具备这种能力，而对微生物来说，这种现象简直是家常便饭。某些情况下，当两个细菌相遇，其中一个会向另一个细菌的方向伸出一根空心管道。当

管道接触到另一个细菌时，一方面，它会通过收缩将两个细菌拉到一起；另一方面，细菌可以通过连接管道向临近的另一个细菌输送自己的DNA。

通过阴茎样的管道向另一个个体传输遗传物质的方式，很容易让人联想到有性生殖。但是细菌的“有性生殖”和人类有着天差地别。它们的交合与人类的不同，不以生殖为目的。基因交换中也不涉及整个基因组的重新洗牌，通常只是交换某几个基因。

细菌还能通过许多其他方式获得新基因。有些细菌可以在别的细胞死亡、破裂或吐纳出内部成分之后吸收外源性的DNA。与其说是阅读，不如说细菌不过是一个在字面意思上“啃书”的傻帽儿，它们除了一把把的纤维素之外什么都得不到，细菌所吞噬的外源性DNA大部分成了食物。只在极偶然的情况下，摄入的外源性DNA会结合到宿主的基因组里并表达出新的蛋白质。

病毒可以用自己的DNA制服比自己大数倍的细胞，插入宿主基因组的病毒DNA重新编程后把活生生的细胞变成了绝望的血汗工厂，成批生产毫无生气的病毒颗粒。在这个过程里，细菌的某些DNA片段会与病毒的基因组融合，使之成为基因转移的载体。这些携带细菌基因片段的新生代病毒离开菌体细胞，将会继续感染下一个倒霉的受害者，通过注入经过融合的遗传物质，将基因从一个细菌传递到另一个细菌。如果人类具有类似的能力，那我们那个高个子篮球选手只需要对着其他人打几个喷嚏，就能把身高的天赋整合到队友的基因组里。

如果所有的基因水平转移都不需要筛选，那么细菌的基因组势必不断扩大直到变得过于臃肿庞杂。过度冗长的DNA链脆弱易断，复制过程会白白浪费许多能量和原料。对大自然来说，浪费是不能容忍的罪过。幸运的是，由于基因融合和删除之间的平衡，过度冗长的基因组不会出现。基因删除是基因错误的副产物，是指细胞在修复和复制DNA的过程中切除错误基因。与每次只涉及一个碱基对的基因突变不同，基因删除往往涉及数千个碱基对和众多基因。只要基因删除没有累及必需基因，细胞就能够继续存活。非致死的基因删除时刻都在发生，它保证了只有有用的基因能够长久留存于基因组内，以及精简的基因组容量。

基因转移与有性生殖的另一个不同点在于，它不仅发生在亲缘关系相近的物种之间，还能够发生在面包酵母与果蝇和微生物与植物之间。尤其在微生物的世界里，哪怕两种微生物的种间差异大如人类和橡树，它们依旧能发生基因转移。这正是基因转移的强大之处，也是它能成就

细菌在代谢进化中的霸主地位的最重要原因。物种之间的差异有多大，它们的代谢方式的差异就有多大。

基因转移通过从一个物种中获得的基因修饰另一个物种，让原本风马牛不相及的优良微生物基因能够融合，正如擅长巴洛克风格和流行唱法的不同微生物终能演绎出一曲风格混搭的乐章。由于不能挑剔或者选择所获得的新基因，而基因的融合随机发生在不同的基因组之间，所以只有部分基因修饰可以改进生物的性状。不过基因转移发生的频率远远超过我们的想象，所以生物进化出新性状的概率其实并不低。即便多数进化的结果乏善可陈，但是宇宙图书馆的书架上摆放了无数本书，在繁多的文字垃圾里依旧有数不清的杰作等待被发掘。

大自然的谱曲能力在人类的朋友大肠杆菌中体现得淋漓尽致，许久以前，科学家曾一度认为大肠杆菌的不同菌株是紧密联系的不同亚种。21世纪初，生物学家首次破解了多种大肠杆菌不同菌株的基因组密码，原本的期望是这些遗传密码高度相似，然而事实却并非如此。有两种大肠杆菌菌株的基因组差异超过了100万个碱基对，相当于它们全部DNA碱基对数量的1/4，意味着每个菌株与另一个菌株有超过1 000个不同的基因。

每过100万年——相当于人类在进化树上与黑猩猩分离至今的1/5，大肠杆菌的基因组就能获得大约60多个新基因，所有的新基因都来自水平转移。它们是基因融合中的成功者，还有很多失败的基因融合没能让细菌留下后代。

如今我们已经掌握了超过1 000种细菌的DNA序列信息，它们证实大肠杆菌菌株间的差异并不是特例，而是普遍规律。细菌基因组的大部分基因都是从别处交换得来的。你可能不会觉得奇怪，不过许多这些基因的起源的确难以追溯。要寻找某个特定基因来源的难度，无异于在国会图书馆中随手拿起一本小说并挑选其中的一小段，然后考证这一段内容在文学史上的影响。1 000多个菌种，甚至1 000种菌种的100倍，也只是由无数种细菌构成的多样性海洋中的区区一滴水而已。更多的细菌甚至还没有被我们发现，而每一种细菌都可能是其他细菌基因的贡献者。

由于细菌基因组中只有大约1/3的基因与代谢有关，所以基因组改变和代谢改变并不总是一一对应的。基因组编码的蛋白质还有许多其他作用，如帮助细菌移动、转运合成所需的物质等。那么如果基因转移主要涉及这些与代谢无关的基因会如何呢？生物进化在代谢图书馆里的步

伐将难以深入，进而导致多数生物的代谢反应高度相似。

实际情况是怎样的呢？几年前，在面对数百种DNA序列已经被阐明的细菌时，我就这样问过自己。这些遗传信息是前人在过去几十年里的研究所得，这项研究发现了数千种独特的酶以及编码它们的基因，让我们能够通过基因辨认相应的酶，并通过酶预测生物具有的生化反应。换句话说，我们可以通过基因组序列预测某种生物的代谢基因型，并对不同生物的代谢基因型进行比较，而这正是我所做的工作。

图3-2以简化的片段对两种生物中的10种酶的代谢基因型作了比较，展示了用这种方式比较代谢基因型的简便性。10种酶中有4种是两种生物都无法合成的，在图中以灰色的0表示，第一种生物编码了其中6种，如你所见，它的基因型数列中有6个1，而第二种生物可以编码其中的5种。

基因型1　0 1 1 0 1 1 1 1 0 0

基因型2　0 1 0 0 1 1 1 1 0 0

距离：D=1/6

图3-2　基因型差异

我们记录了至少被两种生物中的一种所合成的酶的数量（在这里是6），以及只被其中一种生物合成的酶的数量（在这里为1），然后再计算两者的比值（也就是1/6）。如果两者的比值为0，意味着两种生物编码的酶完全相同。如果该比值为1/2，那就意味着其中一种生物合成的酶中有一半能够被另一种生物合成[13]。而如果这个比值为1，那么两种生物中的任何一种能够合成的酶都不会出现在第二种生物体内，两者的代谢差异为可能达到的最大值。这个取值范围为0～1的比值，反映了两种不同的生物在酶学上的差异程度，鉴于这样的描述不太简便，我们姑且以字母D代表这种差异或差距。

如果让你用纸和笔比较数百种细菌的基因型，其中每个基因型都以数千种反应对应的数列编码，其枯燥程度不言而喻，好在忠实可靠的计算机能够在眨眼间完成这些工作。虽然我早就知道细菌的基因组间存在高度差异，但当我要求计算机计算数百对细菌的D值时，我还是被近亲菌株之间的代谢遗传差异震惊了。13种大肠杆菌的不同菌株之间有超过20%的酶互不相同，任意一对微生物之间的平均酶差异达到50%。我还曾经怀疑过是否生活环境相近的细菌，比如都栖息于土壤或都栖息于海

洋，会由于营养条件相近而拥有类似的代谢体系。我又想错了。相似的栖息地并没有能够缩小细菌之间的*D*值差距。

这项工作的结果凸显了自然界在基因重新组合上的惊人尺度。在地球上的每个角落，剧烈的基因拆分和重组都在不断发生。只要是有微生物存在的地方，无论是在海洋深处还是荒凉的山巅，无论是在滚烫的热泉还是寒冷的冰川，无论是在肥沃的平原还是干燥的沙漠，甚至是在我们的体内或体表，生命都在尝试每一种可能的基因新组合，重新解读、重新编译，而后重新布局代谢遗传，片刻也不停歇，造就并不断提升着代谢的多样性。



如果没有读者，一本书就不过是一堆沾染墨水污迹的纤维纸片而已。同样的道理，代谢图书馆里的基因馆藏需要被阅读才能体现它们本身的价值，即每本书所对应的代谢模式应可以代表某种生物可以利用哪些营养物质，又能够合成哪些分子。我们回忆一下某些实实在在、可以被看见的生物表现型，许多代谢表现型如每天的阳光一样朴实可见。比如黑色素，存在于我们体内，可以保护我们的皮肤免受太阳辐射的伤害；存在于在狮子的毛发中，可以帮助这些大猫在狩猎的时候模拟周围的环境；同时它也是章鱼喷出的墨水之所以是黑色的原因。与黑色素类似的分子都是代谢的产物。

其他的色素分子也给树叶、龙虾、花朵以及变色龙染上了相应的颜色，帮助它们防御、求偶，有时甚至根本没有其他的用处。不过代谢表现型并不局限在肤浅的视觉水平，它还存在于我们的眼睛看不见的生化层面，继而不断影响着自然选择。代谢表现型最重要的作用在于保证生物的存活率，归根结底，是与那60多种比色素分子重要得多的基本物质合成有关的能力。存活率，是一种对基因表现型的优劣进行衡量的方式，相当于对一个复杂的故事进行主旨概括，或是一场庭审中的最终判决：如果无法合成所有的基本生命物质，那么就判死刑，并立即执行。任何发生突变致使基本生命物质合成受阻的生物，不是无法存活到可以繁殖的年纪，而是根本无法存活。

为了理解决定生死的表现型，我们必须读懂生物的代谢基因型。这并不容易，不仅是因为基因文本的功能含义要比文本本身复杂得多，我们必须从生物整体上进行把握，考虑不同基因之间的协同效应，还因为

我们的大脑并不擅长解读化学语言。幸运的是，我们可以利用计算机与编程演算，协助我们完成这项工作。

基因型可以告诉我们代谢中涉及哪些催化反应，反应中需要消耗哪些原料分子，又能够合成哪些产物。在解读基因型之前，我们必须首先确定营养物质的来源，俗话说得好，“巧妇难为无米之炊”。然后我们需要检验某种生物的代谢能否利用这种营养物质合成生物必需物质，譬如色氨酸。这对于能够在极端环境中生生不息的生存大师们来说并不难，比如大肠杆菌。这些极端的环境中营养物质稀缺，有时候只有一种糖类可供生物作为能源和碳源。

我们会从环境中存在的营养物质入手，罗列一张清单，枚举所有营养物质通过代谢反应能够获得的产物，然后在生物的基因组内寻找消耗这些产物的代谢反应，并列出这些反应的产物。我们需要重复这几步，直到找到一个或多个反应的产物中包含色氨酸。如果最后没能找到这样的反应，那么这种生物的代谢反应就无法合成色氨酸。

接下来，我们可以把注意力转移到另一种生物基本物质上，可能是另一种氨基酸或是DNA的4种基本单位之一，重复上述整个步骤，以检验每一种构建生物的基本物质是否包含在该物种的代谢反应中。只有能够合成所有生物基本物质的物种才有可能存活。

所有这些工作都是在计算机上完成的，如果使用恰当，计算机运算的速度更快、成本更低，甚至比传统的实验结果更可靠。但正如纸上谈兵并不等于可以攻城略地，对生物学家来说，任何没有经过实验验证的计算结果都需要谨慎对待。正如工厂会对出产的产品进行随机抽查，我们也需要抽选一种已知基因型的生物，将其培养在成分已知的环境中，然后静观其变。其实也可以说是冷眼旁观，任它们自生自灭。这种工作早就已经有人做过了，他们实验的对象包括了数百种大肠杆菌的变种菌株，这些变种大肠杆菌都通过基因工程敲除了某一种酶。实验结果与计算机演算结果高度吻合：超过90%的菌株实验与演算结果相符。

大多数知道这项演算实验的生物学家都把它当作理所当然，并不觉得这项工作有多稀奇。但事实上，这远不止是稀奇而已，能够通过计算机预测生物生存能力的技术具有深远的开拓性意义，它是数百年的传统生物学研究与现代计算机科学结合的产物。达尔文以及在他之后的几代生物学家大概做梦也想不到，有朝一日世界上会出现这样的技术，而计算机技术对于我们理解代谢进化，理解大自然如何创造出了新的代谢模式至关重要。

对于任何已经了解其代谢功能的生物而言，在任何成分已知的环境中，无论是极地土壤、热带雨林、海底深渊，抑或是山地草甸，我们都可以用这种算法进行模拟。这种算法同样适用于评估代谢表现型的任何层面，比如预测代谢反应中能够合成的所有分子。不过，在能够进行演算的所有方面中，合成生物基本物质的新手段与利用能源物质的广泛适应性是最重要的层面，而生物存活率则是这一切的根本意义所在。新的代谢能力是不断驱动生命拓展最前沿阵地的引擎。

利用新物质作为燃料的能力之所以如此重要，其原因非常简单：无论一种代谢方式在今天看来有多成功，由于世界的瞬息万变，它几乎注定会在未来的某一天掉下神坛，正如将随着不可再生的化石燃料日渐枯竭而凋零的全球经济。环境中的化学成分也是一样，营养物质总是旧去新来，从来不会一成不变。依赖某几种特定营养物质的生物容易走入进化上的死胡同。生命如果想繁衍下去，就必须寻求新的代谢方式。万幸的是，许多不同种类的分子都可以为生命体提供能量和必需的化学元素，有我们熟悉的葡萄糖和蔗糖分子，也有一些可能相对陌生，比如剧毒的五氯苯酚。

只需要较少的几种原料分子，就可以组合出数量惊人的代谢类型。它们的可能数量相当巨大，不过并不是所有这些代谢表现型都能保证生物的存活。如果想对这个计算题有个大致的印象，我们来看图3-3中列出的100种潜在的燃料物质。然后，我们来统计一下某种你感兴趣的动物、植物或细菌是否能够利用某种特定的物质，比如葡萄糖。如果这种生物可以利用葡萄糖合成所有其他所需的基本物质，就把葡萄糖标记为“1”，否则标记为“0”。接着，我们对下一种物质重复同样的步骤，直到所有物质旁边都有相应的“0”或“1”标记。这个清单中的每一个“1”都意味着你考量的生物能够只利用对应的物质合成所有必需的基本物质。

燃料物质	表现型
葡萄糖	1
乙醇	0
•	•
•	•
•	•
蔗糖	0
果糖	1
•	•
•	•
•	•
柠檬酸盐	1
醋酸	1

图3-3　代谢表现型

完成编码后得到的“0”“1”数列描述了给定的新陈代谢利用不同燃料分子维系生命的能力。这是表达一种生物代谢表现型的精简方式。像大肠杆菌这样的代谢通常能够依靠数十种不同的碳源生存下去，因此它们的表现型数列中有很多“1”。与之相对，某些精专的生存大师只能利用为数不多的碳源，所以它们的表现型数列里多数都是“0”。

在计算100种能源物质能够组合出多少种代谢表现型前，我们只需要牢记，对于每一种物质而言，生物体只有能够或不能够依靠这种物质生存两种结果，除此之外没有第三种可能，因此所有的可能代谢表现型是100个2相乘，也就是2^{100}。这个数量超过了10^{30}，也就是1后面跟着30个零，虽然和现实中实际存在的可能表现型数量相比还有差距，但已经是一个天文数字了，因为这数字已经比银河系中的恒星数量要多了，如果我们非要拿来比较的话，后者仅为10^{11}，也就是“区区”1 000亿。

现在你可能意识到了：我在上一章就提到过，现代综合进化论的缺陷是它过于忽视生物高度复杂的表现型。现在看来这可不是我在开玩笑。

表现型的巨大数量同时也意味着代谢进化的巨大潜力。图3-4中给出了一个例子。图中左侧展示了某种代谢表现型能够利用的碳源，但是这种代谢方式无法利用乙醇，因此在乙醇旁标记为“0”。无论是不是通过基因转移获得的，一个新的基因可以通过改变基因型进而让表现型具有代谢乙醇的能力。如果该变异使代谢乙醇成为可能，我们就把“0”改为“1”。由于每一种新出现的代谢表现型都可以用这种标记方式表示：通过把代谢表现型中的某个“0”改成“1”，所以理论上来说，代谢表现型

的数量越多，生物的进化潜力就越大。

葡萄糖	1		1
乙醇	0		1
·	·		·
·	·		·
·	·		·
蔗糖	0	→	0
果糖	1		1
·	·		·
·	·		·
·	·		·
柠檬酸盐	1		1
醋酸	1		1

图3-4　代谢进化



由于代谢类型的数量巨大，远远超过宇宙中的氢原子数，所以要腾出一块地方，专门建一栋收纳所有表现型文本的图书馆显得异常艰巨。此外，如果要在这个图书馆里迅速检索到某册馆藏，那么馆内的收藏必须高度有序。我的办公室里有个小图书室，我在那儿只要几秒钟就可以找到以前买的那本《物种起源》，作者正是达尔文。不过，如果要在一个常规大小的大学图书馆里边晃悠边找某本特定的书可就没那么简单了。而如果《物种起源》被人放错了书架，那么可能就永远消失在这个图书馆里了。同样的错误在一所超宇宙数量级的图书馆里导致的后果只会更糟糕。宇宙图书馆里很可能藏着解开长生不老之谜的秘籍，就算没有，也肯定有配方教你如何煮出完美的火鸡填料。但由于图书馆实在是太大了，如果我们不知道这些书摆在哪里，那么我们可能永远也找不到。

一种相当简便的图书馆归档方式是把书按照内容的相关程度摆放。人类的图书管理员在归类不同印刷版次的同一本书时就会用这种方式。如果代谢图书馆在归类书籍的时候也遵循相同的原则，那么越相似的文本之间应该距离越近。但在讨论归档之前我们首先要解决一个问题：采购或者制作这个图书馆需要的书架将是一件痛苦的活计。

在现实的图书馆里，每本书都与另外两本书相邻，左右各一本，即使算上书架上下的书，那么一本书最多也只与四本书相邻。但代谢图书馆里的每本书会与多少本其他的书相邻呢？这里我们可以回忆一下代谢

图书馆里那些每本由5 000个字母组成的馆藏。每本相邻的书都只相差一个字母，相邻的代谢基因型之间只差一个生化反应。（两个代谢基因型之间的差异无法比一个更小，而当两者差距进一步拉大时，它们就不会被相邻摆放了。）

我们假设，在与代谢图书馆中任何一本书相邻的其他书中，第一本与原书的第一个字母与原书不同，第二本则是第二个字母不同，每一本相邻的书都依次与原书对应的字母不同，直到最后一个字母。换句话说，代谢图书馆里的每本馆藏不是与两本，也不是与四本，而是与上千本书相邻，具体的数目取决于生化反应数量的多少，相邻的馆藏之间只相差一个字母，也就是一个生化反应。能够满足如此陈列要求的书架可不是那么容易找到的。

为了帮助你理解这种情况有多复杂，我们先从更简单的情况开始讨论，最简单的化学世界莫过于只有一种化学反应。在那个世界的代谢图书馆里只有两本馆藏。一本的内容是“1”，由唯一的一种化学反应构成；而另一本是“0”，代表该种代谢类型不具有该反应。图3-5中a图的两个端点和连接两者的直线就代表这种情况。

比直线稍微丰满一点的世界由两个化学反应构成，相应的代谢图书馆规模将扩建到4（2^2）种可能的馆藏。其中之一同时拥有两种反应（11），有两种代谢型拥有两个反应的其中一个（10，01），第四种代谢型则同时缺乏两种反应（00）。如图3-5的b图所示，这种情况下，每个代谢基因型就如同一个正方形的四个顶点。

可能你已经明白我接下来要说的事了。下一个级别的世界里包含了三种化学反应以及8（2^3）种可能的代谢类型，我们用一个立方体的顶点表示这8种代谢。而在一个包含4种化学反应的世界里，我们能够得到16（2^4）种可能的代谢型，但是哪种几何图形能够与之对应呢？随着例子中化学反应的数量从一到二再到三，对应的代谢型分别占据了一条直线、一个正方形和一个立方体的顶点，不同的几何图形又分别对应一维、二维和三维空间。尽管四维或者更高维度的空间很难用视觉图形的方式呈现，但和它们打交道依旧是数学家们的家常便饭，因为他们能够将已有的几何规则演绎到这些多维空间中。

就像四边形和立方体，我们所寻找的几何图形的每条边长都应当相同，不同的边相交需要形成一个恰当的角度。如此我们便能够找到一个四维的超立方体。图3-5中的d图就以几何技法展示了超几何体在平面上的视觉效果。具有四维空间的超立方体有16个顶点，每一个顶点对应一

种代谢类型，即从0000到1111，不过我们并没有在图中一一标记出。

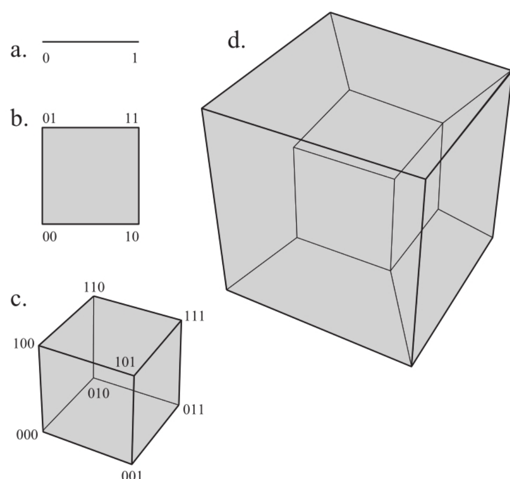


图3-5　超立方体

这种绘图的方式在面对五维空间时就显得力不从心了，遑论更高维的空间。不过虽然把高维空间的图形视觉化有点不切实际，但是它们依旧遵循与三维空间的图形一样的原则：边等长、恰当的角度以及与每一种代谢型相对应的顶点。符合这些原则的高维度几何图形，其性质恰好符合代谢图书馆的需要。

一个正方形有4个顶点，在立方体中这个数字会翻倍到8，而在一个四维的超立方体中顶点数量会变为16。空间每增加一个维度，对应几何体的顶点数量就翻一倍。当我们讨论五千维空间的时候，顶点数量就达到了$2^{5\,000}$个，也就是代谢图书馆的规模。换句话说，我们可以把这些馆藏摆放在一个五千维空间里的超几何体的顶点上。面对五千维空间，身处卑微三维空间的我们几乎束手无策，这就是为什么代谢图书馆里不能用现成的普通书架。它需要一个五千维的解决方案才能维持馆内的运营。

除了馆藏的摆放问题，超几何体还可以很好地解决馆藏之间的相邻问题。在相对简单的三维空间里，每一本图书馆的馆藏，也就是立方体的每个顶点，都与另外三个顶点相连。我们以其中一个顶点为例，比如图3-5　c图中的数列100，你可以沿着从该点伸出的边到达与100相邻的顶点。与之相邻的顶点要么比100多出第三种反应，对应的代谢编号为101；要么比100多出第二种反应，对应的代谢编号为110；或者缺乏第一种反应，对应的代谢编号为000。所有相邻的顶点：101、110以及

000，都与100仅相差一位数字。立方体中任何一个顶点的情况都与例子中的这个顶点一样：它们都与三个其他顶点相关联。

类似地，五千维的超几何体中，每一种代谢型都和与维度数一样多的其他顶点相邻，也就是5 000个。从每个代谢型所在的顶点出发，你有5 000个方向可以选择，只需要跨出一步，你就可以到达5 000个与之关联的顶点中的一个，而且相邻的代谢型都只相差一种化学反应。要么多一种，这种情况下某一个编号中的0就是另一个相邻编号中的1；要么少一种，也就是某个1变成0。

生物进化的过程就像参观代谢图书馆，基因删除和基因转移就是生命在图书馆里移动的方式，让它们从一本馆藏跳到下一本，而通常就是相邻的那一本。每本书相邻的所有其他书可以被称为一个“社区”（neighborhood），对于生物进化来说，这个社区如同现实生活中真实的城市社区，对人们的生活而言，具有同等的重要性。城市社区的有用之处体现在它的便捷性上：人们需要的东西都在几步之遥，代谢图书馆的“社区”也是一样的道理。进化只需要对基因型进行微不足道的一点修改，就可以搬进自己邻居的家里。不过城市社区里的居民只能沿着东、南、西、北四个基本的方向行走，而进化有5 000个不同的方向可去。（这个复杂的场面你最好连想都不要去想。）因此，一种代谢型身处的社区肯定比你所在的小区有趣且丰富得多。我们很快就将看到，代谢图书馆惊人的多样性在进化的创造性中具有的重要性。

随着时间的推移，某种生物基因组中积累的改变越来越多，它也渐行渐远，进而到达图书馆内距离更远的书架。为了估算距离，我们需要寻找一种度量的手段。没有度量的能力，我们就无从得知进化如何周旋于不同的书架之间，图书馆就像一个迷宫，我们将迷失在毫无意义的书堆之间。幸运的是，我在研究中所用的基因型差距值*D*可以胜任度量的工作。*D*值能够代表图书馆中两个代谢文本之间的距离大小，事实上，它已经告诉我们某些生物的代谢型相距甚远。除此之外，它为我们提供的另一个洞见才是重点：代谢进化能够在代谢图书馆中穿越惊人的距离，而许多进化的文本不管披着何种外衣，它们诉说的故事寓意都是相似的。



终有一天我们将能够破译数以百万计的代谢文本，但是对于超宇宙

数量级的代谢图书馆来说，这也不过是沧海一粟，甚至仅仅是宇宙中的几粒尘埃而已，代谢图书馆里的馆藏远远超过地球上所有曾经存在过的生命的总和。尽管已经经历了**38**亿年的进化，生命依旧只是徘徊在图书馆的某个角落。

在生物进化的数十亿年间，大自然完全不需要顾虑会在宇宙图书馆的下一个拐角遇到什么样的新馆藏。但是如果人类希望理解图书馆，而不是在其中漫无目的地游荡，我们就要学会在图书馆里寻找那些有意义的生命文本。不仅如此，我们还要学会对已知的文本进行分类，如同杜威十进制图书分类法[14]或是美国国会图书馆分类法[15]那样，先按照不同的主题进行归类，如艺术史、经济学、语言学……然后再以更小的类别细分，比如语言学中还可以分为罗曼语、德语、斯拉夫语等。代谢的表现型，也就是代谢基因文本的具体含义，是代谢图书馆天然具有的分类方式。代谢图书馆里的馆藏比现实图书馆中的书要多得多，不过这仅仅是因为代谢图书馆本身的规模过于庞大。

分类法就如同一张探索代谢图书馆的地图，我们如果想要某种表现型，那么一张基因型-表现型地图可以指引我们去哪里寻找它的基因型。如果没有这张地图，我们就无从得知题材类似的馆藏是摆放在一起还是散落于图书馆内各处，虽然在人类的图书馆里它们总是被安排在一起；我们也不知道同一个书架上是否会陈列主题不同的作品，凡此种种。由于没有图书管理员，所以我们需要像古埃及时期游历世界绘制大陆形状的航海家们一样，通过自己在图书馆里游荡和探索亲手绘制这幅地图。代谢图书馆巨大的规模使得我们几乎不可能摸清它的每一本馆藏，不过我们依旧可以描绘大陆、山川、河流、湖泊以及沙漠的轮廓，以期能够从模糊的形状里窥得壮美山河的蛛丝马迹。

但是该从哪里着手，又该沿着哪里探索呢？

首先，我们需要找一片拼图来为我们指路。以任何一种代谢表现型为例，比如依靠葡萄糖存活的代谢性状，假设如果代谢图书馆内超过10^{1500}个代谢文本中只有一种能够表达这种性状会怎么样？如今地球上的细菌总数大约是5×10^{30}个，这个数量十分巨大，1后面要跟着30个0。我们可以假设自从生命出现起，每一个细菌以一秒一种的速度尝试新的代谢模式，那么在已经过去的将近40亿年里，它们总共只尝试了大约10^{48}种代谢模式。细菌们随机找到那种仅有的、能够利用葡萄糖进行代谢的概率微乎其微，还不到$1/10^{1450}$。这个概率小得几乎没有任何实际意义。换句话说，这种盲目的搜寻方式最终将无法令细菌获得相应的性

状。

一方面，寻找到某种特定性状的概率是渺茫的；另一方面，生命表现的多样性表明，进化寻找新性状的能力无须置疑。这也意味着上述假想的情况是错误的。代谢图书馆中包含葡萄糖的文本肯定不止一本，很可能有许多能够利用葡萄糖的不同代谢模式。

为了找到这些代谢文本，我们来模仿一下进化曾经做过的事：尝试探索图书馆和编辑基因组，也就是对基因组进行一系列删除和转移，消除或增加某些基因、酶及生化反应。从哪里开始入手其实并不重要，我们可以选择代谢图书馆里任何一个馆藏，选择任何一本包含葡萄糖或者其他能源物质代谢的文本。

现在我们从一个包含葡萄糖代谢的代谢文本开始，随机删除一个文本中已有的反应，或者向文本中加入一个已知的生化反应。这个经过修改的文本在面对大自然的审阅时，得到的回复往往简单粗暴：生或者死。但是作为科学家，我们不用如此循规蹈矩。我们能够通过算法解读代谢文本的含义，如果结果显示新的文本所代表的代谢类型不能利用葡萄糖维持存活，那么就返回原始文本，重新删除或添加一个生化反应。不要忘记，可选的目标基因有5 000种之多。只要经过修改的代谢依旧能够利用葡萄糖，那么这种修改就可以继续下去，接着添加或删除第二个基因，演算对应的表现型，再评估，如此循环往复。

也就是说，我们从某个起点开始，首先到达与它相邻的文本，再到相邻文本的相邻文本，而后再到相邻文本的相邻文本的相邻文本，直到我们弄清楚在不改变代谢表现型的前提下，即对葡萄糖的利用能力，能够到达的极限距离。由于文本的每一次改变都是随机的，所以在代谢图书馆中的这种移动是一种随机游走（random walk），就像一个走出酒吧的醉汉跌跌撞撞地在路上晃荡，碰巧撞进了自己的家门，只是有一点不同：在代谢图书馆里的每一步都必须踩在主题相同的文本上，也就是沿着相同的表现型前进。

如果只有一种代谢类型里包含葡萄糖代谢，由于它没有所属的“社区”，那么通过随机游走无法到达任何地方，我们只会停在原地止步不前。不过即便与某种表现型对应的文本不止一本，但只要它们散落在图书馆内不同的角落而没有相邻，我们同样无法从其中一本馆藏出发，直接跨越其他代谢抵达同类型文本。哪怕这些零星的文本聚集在一处，随机游走的脚步也不一定会走得更远。作为起点的文本可能还有几个不多的邻居，但是这些邻里未必有着它们自己的邻居。

只有当同一主题的文本具有相当的数量时，我们才能循着它们探索整个图书馆。不过如此一来我们又将面对一个新的问题：计算量。计算一个代谢文本的含义不算什么大事，但是如果要分析随机游走过程中的数千个文本就没那么容易了，更不要说随机游走中前进的每一步都有数千种不同的可能性。普通的家用计算机大概需要数年甚至数十年才能完成类似的计算。利用互联网连接的一个计算机集群能够令我们获得更高的计算速度，但其巨额费用也令人难以承担。

在度过漫长的攻读博士学位的阶段后，我成了一名博士后并最终在美国的一所研究型大学被授予终身教授职位，在这个过程中，资助进化研究的经费日渐枯竭、每况愈下。研究经费的匮乏正好与我远在欧洲的家人的患病撞车，所以那一年，当一份来自瑞士的工作邀请摆在我面前时，我其实早已做好了跨过大西洋、回到故乡的准备。

一直以来我都知道，瑞士是世界科学的引领者之一，瑞士科学研究的产量惊人、水平拔尖。而科研成功的背后离不开其世界一流的公共教育体系，对学术研究的慷慨资助以及宜人的居住环境。我为不得不离开在美国一起研究学术的同事们而觉得遗憾，同时又对能够进入瑞士的科研圈感到诚惶诚恐。而最重要的是，这份工作不仅能够从资金上支持一个计算机集群，同时也足以维持运营一个现代化的实验室。更妙的是，我能够在世界范围内招募不少和我有类似想法的研究同行。这份工作让我不敢再有别的奢求。



2006年一个天高气爽的秋日，我坐在自己的办公室里，它位于苏黎世大学一栋外形简朴优雅的建筑内，外墙的玻璃和金属在阳光下反射着微光，勾勒出大楼的轮廓。那天，一个年轻的葡萄牙人走了进来，他长相英俊、说话轻声细语，深褐色的眼睛里闪着好奇的光芒，他礼貌地笑了笑之后，告诉我说他叫若昂·罗德里格斯（Joao Rodrigues）。

若昂一直在研究物理学，同时也发现生物学中有许多亟待解决的问题。他在寻找一个新的挑战，希望能够通过打破两个学科之间的壁垒获得自己的博士学位。若昂对生物学的了解有限，但他具备很多生物学家没有的本事：他十分擅长数学和计算机编程，也亲自操刀过许多大型和复杂的演算项目。在浏览他简历的时候，我简直抑制不住内心的狂喜。若昂拥有的能力恰好是探索代谢图书馆所必需的。在对他的面试中，我

情不自禁，与他分享了我对自然进化的见解。幸运的是，我们一拍即合。我看到他的眼神里闪着光。最终，若昂欣然接受了这份工作。

在我的实验室里，若昂的学术背景不仅不算特殊，反而是个中典型。这里的研究人员来自十多个不同的国家，有美国以及欧洲、亚洲和澳洲的国家，他们的研究领域各异，包括生物学、化学、物理学和数学。这一切都是刻意而为，因为我们所面对的难题需要不同领域协同合作，因此我喜欢把我们的工作比作进化本身：研究也需要与时俱进，将各种传统研究方式进行重新组合——不是酶的组合，而是学术技能，这将大有助益。

我们搭建的由100台电脑组成的计算机集群依旧不能使我放心，我担心其计算速度依然不足以让我们离开代谢图书馆里的第一个书架，不过若昂的计算机技能像魔法一样让我印象深刻。他巧妙地提高了计算机的工作效率，使得它们的计算能力提高了数倍，最终把我们远远地送到了图书馆深处。

若昂的演算始于一种广为所知的代谢：大肠杆菌代谢葡萄糖，它能够以这种单糖为原料合成所有必需的60种基本物质。为了验证大肠杆菌的代谢方式是否独一无二，若昂首先设计了1 000多种大肠杆菌的“邻里代谢”，它们中的每一种都与大肠杆菌的代谢相差仅一个生化反应。如果大肠杆菌的代谢是一本利用葡萄糖合成所有必需物质的说明书，那么这些设计出来的代谢就是这本说明书最接近的山寨版本。首要的问题在于：这些人造的代谢模式是否具备所有利用葡萄糖合成必需物质的信息？

经过演算若昂很快发现，不是一个、两个或者三个，而是数百个大肠杆菌的“邻里代谢”能够依靠葡萄糖维系生命。这个发现说明了一个简单而重要的结论：认为大肠杆菌代谢葡萄糖这个性状独一无二是一种错误的偏见，基因型所在的社区里包含众多类似的基因型。但是更让我们吃惊的发现还在后面。

若昂利用大肠杆菌作为起点深入代谢图书馆的探险，把他引向了距离这个起点越来越远的地方。演算的目的是为了测试我们能够以这种方式深入到图书馆的何处：我们希望从一种能够支持生命存活的代谢到达与它相邻的代谢，再从相邻的代谢到与之相邻的代谢，依此类推，同时保持生物利用葡萄糖的能力。在保证主题不变的基础上，代谢基因型能够被编辑的最大限度是多少呢？当若昂给我展示计算结果时，我的第一反应是不相信。他找到的距离最远的代谢，也就是*D*值最高的代谢，与

我们开始时的大肠杆菌代谢仅有20%的相似性。从算法的角度来看，我们几乎已经穿过了整个图书馆，那可是相距最远的馆藏之间80%的距离。在这个距离上，只要再深入一步，我们就找不到包含葡萄糖代谢的文本了。

因为担心单一的实验可能不具有代表性，我要求若昂再多执行一些随机游走的计算，一共1 000个。对每一个基因型的运算都以保证相同的代谢为前提，演算可能到达的最远距离（*D*值），对相邻基因型进行尽可能多的尝试。这不是不可能的，因为这座图书馆里最不缺的就是岔路。当结果返还到我手中后，我又一次被震惊了。所有的随机游走都到达了和第一个结果几乎一样远的位置。每一个结果与原始大肠杆菌的代谢模式的差距都几乎达到了80%。

我们实验室的研究员找到了1 000多种与大肠杆菌代谢基因相差巨大的代谢模式，它们唯一的共同点只有能够利用葡萄糖作为单一碳源和能源进行合成代谢这一点。如果我们继续下去，肯定还能找到更多类似的代谢型，多到我们数不过来。不过我们后来终于学会了如何估计某些代谢型在图书馆内的数量，比如包含2 000个代谢反应、能够利用葡萄糖的代谢文本大约有10^{750}种。

不要说图书馆，哪怕是仅仅包含葡萄糖代谢的文本就已经是一个超宇宙常数了。代谢图书馆里堆到天花板的那些书，其实不过是在用不同的方式诉说着同一个故事。

万万没想到的是，我们在探索过程中还发现了这座图书馆的一个更诡秘的特征。那数千个随机游走的算法并没有终结在文本内容相同的书堆中，也就是一小群类似的代谢反应模式里。随机游走沿途经过的所有代谢模式，不论是与原本的大肠杆菌还是其他模式相比，都一样天差地别。每种代谢基因型所编码的代谢模式，包含的生化反应都各不相同。不像现实中的图书馆会设置历史书籍区或科学书籍区，代谢图书馆并没有严格地按代谢的类别划分区域。

最让我们惊讶的是，当我们以任意一种代谢模式作为新起点，以保证生物的存活为前提，以保证某种特定的性状不变为前提进行随机游走时，我们最终总是能找到一些类似的文本，而不论它们离起点有多远。这似乎意味着，图书馆中主题相同的馆藏相互联系，形成了一张网络，我把这张网络称为基因型网络（genotype network）。它看起来可能有点像图3-6中那张由直线构成的网络，整个矩形即代表代谢图书馆，而其中的线段将同一个社区的文本（图中的圆圈）连在一起。这张图只能作

为视觉上的辅助，以二维代替五千维，以有限的圆圈代表难以计数的文本，不过除此之外，我们暂时还没有更好的办法来演示如此诡异的图书馆。

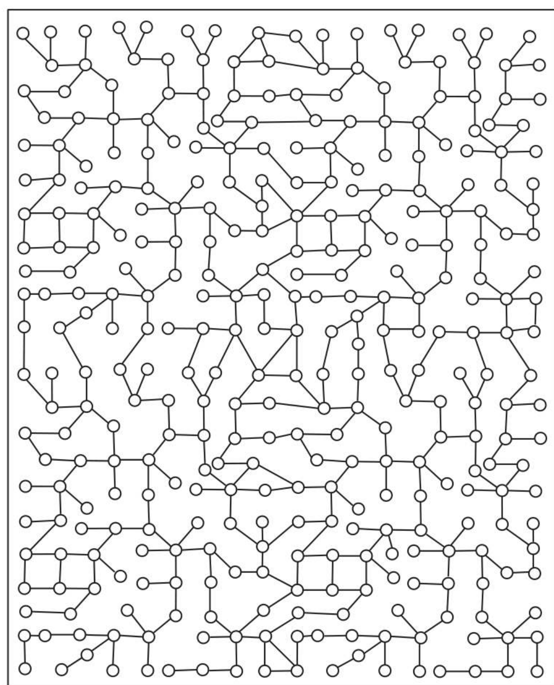


图3-6　基因型网络示意图

在一座普通的图书馆里，你很有可能在历史图书区找到一本有关查尔斯·达尔文的书，当然你也可以在传记图书区找到类似的书。而如果在一所采用美国国会图书馆分类法的大型学术图书馆里，你应该能够在QH区（代表“科学：自然史，生物学”）、DA区（“世界史，不列颠”）、GN区（“人类学”）、PR区（“英国文学”），甚至BL区（“宗教，神话，理性主义”）找到类似的书。但是在代谢图书馆的组织原则里，你找不到一丝这样分区的痕迹，你也找不出两本分别在HM区（“社会学，概论”）和BT区（“教化论”）的书有何关联，除非你沿着“达尔文生平”这个线索，循着一本又一本相邻的书在图书馆内前进。没有这些以不同口吻和角度描写达尔文的馆藏作为线索，你很快就会迷失在浩瀚的书海里，寸步难行。

我们在代谢图书馆里正是这么做的。含义相同的无数文本在图书馆内就像散布在宇宙中的星辰，中间隔着广袤的未知空间。但实际上它们并非处于孤立状态。它们之间以城际高速相连，高速路上灯火通明。

到这里为止，我们仅仅对一个主题的馆藏完成了分类，即以葡萄糖为维生物质的代谢，除此之外还有许多其他的主题。有的代谢类型能够以乙醇、乙酸以及数十种其他物质支持生命。我们以相同的归档方式对它们进行了制图：以某种代谢表现型作为前提进行随机游走的演算，如以能够利用乙醇为例，直到我们无法在保有这种性状的基础上再前进一步。我们针对80种不同的物质进行了计算，而每次我们都能看到类似的模式。建立在同一种物质代谢基础上的基因型的代谢相似度可以仅为**20%**，正是它们在代谢图书馆中连成了一张宽广而稀疏的网。

有了这个普遍规律作为初步结果，我们便斗胆将目标转向了能够同时利用多种物质维生的代谢类型，如能够同时利用乙醇、葡萄糖以及乙酸。（能够利用多种物质的优势显而易见：生物不至于因为其中一种物质耗尽而无法生存。）由于这种代谢方式更复杂，所以会不会只有图书馆内某个角落里为数不多的代谢型能够实现？事实证明我们又想错了。我们研究了能够利用5种、10种、20种乃至60种不同物质的代谢型。每一个随机游走的演算都到达了距离起点相当远的位置。即便是同样能够利用60种不同物质的不同代谢模式之间也只有30%的生化反应相同。即使如此，这些数量在万亿级别的表现型相同的代谢，依旧组成了一张相互联系的基因型网络。

到了这一步，我几近狂喜。我们偶然发现了代谢图书馆内组织构建的最基本原则。首先，许多代谢型都能够以相同的物质作为能源，这与具体的能源物质种类关系不大。生物通过对不同化学反应进行千奇百怪的组合，合成了必需物质。其次，相同的能源并不意味着相同的代谢，这些代谢型往往只有一小部分生化反应存在交集。最后，我们演算得到的代谢型都在一张巨大的网络中相互联系，这张网络就是基因型网络。每一类代谢都有各自的基因型网络，所有的网络在代谢图书馆里互相纵横交织，仿佛一块致密的绸缎。

我们以有限的预算完成了这项工作，我们的计算机集群在面对代谢图书馆里的馆藏数量时依旧显得力不从心。不过我们依旧为这个大得超出想象的图书馆绘制了一张粗糙的地图，仿佛乘坐浴缸进行了一次环球旅行。



同一表现型的不同代谢文本提高了我们找到该性状的概率，而且是

成倍提高。此外，进化可不仅仅是一名在图书馆里闲逛的读者。相反，它会招募大量生物进入这座图书馆里寻找新的文本，每发生一次基因转移，生物在代谢图书馆中就深入一步。有着数十亿读者在朝着图书馆内不同的方向展开探索。

与我们在现实生活中逛图书馆相比，进化探索自然图书馆的方式还有一个不同之处。为了便于理解，我们可以假设有一个生物个体遭遇了一场变故，很可能是一个基因的删除，因此从安全前行的道路上偏离，与原本维持它生存的代谢文本失之交臂。发生在它身上的基因删除可能会摧毁某个关键分子的合成能力，而这个个体毫无生还的可能，自然选择将慷慨地赐予它死亡。这就是代谢图书馆，在那里，有些读者会在延绵数代的探索中消亡，而有的则得以生还。

从局外人的角度来看，图书馆中的探索者们，无论是细菌还是蓝鲸，并没有比尘埃泥土特别到哪里去。在自然图书馆面前，生命卑微得像无根的野草，在世间到处漂泊流浪。无数生命用自己的身躯试验着不同的化学反应组合，不断地试验，不断地重复。有些一命呜呼，有些则侥幸生还，继而把自己的经验传递给下一代。生命犹如风中翻腾的黄沙，生命进化的过程并不比无处安身的风尘高贵多少。

基因型网络就是那股风，没有它，生命的黄沙就失去了前行的动力。如果代谢某种物质的解决方案是唯一的，那么所有探索图书馆的读者就不得不挤在某本书周围。任何企图到附近书架开小差的个体都会被淘汰。而如果内容类似的文本稍微多一些，读者们也只能围在图书馆的某一小块区域内。多亏基因型网络的存在，生命才能在保证原有性状的同时，深入探索图书馆的各个角落。

生物进化的关键因素有两个，基因型网络只是其中之一。我们现在来看看第二个因素：代谢图书馆中社区内性状的高度多样性。

想象一下，一小块泥土中有着数十亿个细菌，只要偶尔给它们一点接济，比如一片掉落的叶子，一具腐烂的尸骸，或者一个从树上掉下的熟透的苹果，百无禁忌，它们就能生生不息。这些食物中营养物质丰沛，不过前提必须是细菌有能耐消化和吸收它们。换句话说，也就是细菌有适当的酶，可以利用外来的物质合成自己需要的生物成分。当可用的食物全部耗尽，只要有一个细菌拥有利用不同物质的能力，它就很可能会成为其他嗷嗷待哺的细菌的救世主。此时，新的性状就是微生物们延续生命的关键。

如果我们考量100种利用不同物质的代谢，它们之间的相互组合方式也将轻松超过10^{30}种，而上述那位救世主的代谢模式只是这么多组合里的其中一种而已。要把这10^{30}种代谢放在图书馆里的一个社区内自然是不现实的。每个社区大约只够容纳100多个不同的文本，这仅仅相当于所有代谢表现型的$1/10^{26}$。这就好比你随手从纽约公共图书馆里借走几本书用来填补你空空如也的床头柜，然后希望这其中包含了达尔文的《物种起源》。换句话说，你在白日做梦。但是，如果是一群读者循着某种指引分散深入到图书馆内，那么这个概率就大不相同了。由于基因型网络巨大无比，所以这群读者能够由此接触到数千个社区，这将大大提高找到目标文本的可能性。

为了验证这种组织形式是否真的存在，我们挑选了成对的代谢文本，每对文本的表现型都相同（比如利用葡萄糖的能力），除此之外，两个文本没有其他共同点。我们把两个代谢文本编码为A和B，它们位于自然图书馆内两个不同的位置，即它们包含的生化反应几乎各不相同，但它们又都是同一个基因型网络内的成员。接下来我们来检查它们所在社区里的5 000多个其他文本，其中的某些文本也同样具有利用葡萄糖的能力，也就是和我们挑选的文本属于同一张基因型网络，但也有一些失去了某些关键的生化反应，最终导致生物死亡。还有一些相邻文本——它们一直是我们关注的焦点，赋予了生物利用新物质的能力，比如利用乙醇或果糖。

对于这些基因型网络我们想问的是：A文本所在社区中的文本，即那些与A文本只相差一个生化反应的代谢模式，是否与B文本所在社区内的文本不同？如果A文本的相邻文本中包含能够利用乙醇和果糖的代谢方式，那么B文本所在的社区里会不会也有能够利用其他物质的代谢，比如，利用醋酸和蔗糖？

在分析了数千对代谢文本以及它们的表现型之后，我们发现之前的预设是正确的。文本所在的社区内往往有着控制新性状的文本，而不同社区内文本的表现型也十分不同。许多代谢性状都是某个社区所特有的，不会出现在其他社区中。（这是因为每种表现型都有自己所在的基因型网络，同时也意味着不同的基因型网络相互交织的方式极其复杂。）

在计算机的帮助下，我们进行了更进一步的探索。我们再一次漫步于代谢图书馆的基因型网络中，只是这次我们担任的是拿着笔记本的仓库管理员的职责，我们想要把所有与沿途文本直接相邻的文本记录下

来，而这些新的文本是最触手可及的。在开始前进之前，我们列出了所有起点文本附近的新性状，之后我们跨出第一步，继续检查当前所有的相邻文本。如果新社区内包含原先没有的性状，我们就把它们加到列表里，然后再往前走一步，检查新的社区，加上新的性状，如此反复，直到走出数千步。我们已经知道不同的社区中包含的性状往往不同，所以我们猜测，随着愈发深入图书馆，列表上记录的新性状会越多，但是我们迟早会记录完所有的性状。

事实证明，我们的想法大错特错。记事本马上就写满了，但新的性状还是源源不断地涌现出来。

为了排除研究的偶然性，我们继而重复了很多次类似的尝试，从不同的起点开始，分析依靠不同物质作为能源的代谢方式。我们还增加了实验的样本，想要计算出它们到底能找到多少新的性状。在每一次演算里，新的性状总是稳步增加，毫无衰减和停歇的迹象。不管我们的演算持续多久，无论是100步，1 000步还是10 000步，也无论是一小时，一天还是一周，直到我们用尽时间，或者有新的工作要做。最终，我们意识到，在有生之年我们恐怕是看不到代谢进化江郎才尽的那一天了。

代谢图书馆里的新性状几乎取之不尽。基因型网络和社区多样性亦然，它们是进化发生的两个关键。基因型网络确保了生物探索自然图书馆的能力，没有基因型网络，生物一不小心就会踏入万劫不复的境地。而如果没有社区，沿着基因型网络进行的探索就失去了意义：网络中的性状都一样，对其中某个性状的探索不会带来任何新的性状。

人类图书管理员在管理现实的图书馆时可没有这样的本事。且不说去哪里找用数千种不同的方式讲述同一个故事的书，即便有，也没有图书管理员会模仿自然图书馆的组织形式，在一个主题区域里摆放内容不同的各种书籍，他们也无法把含义不同的书安排在主题相似的文本附近。

不过只要仔细思考就会发现，代谢图书馆并不是什么疯子脑袋里的奇怪想法。人类的图书馆之所以非常实用，仅仅是因为图书管理员按照我们的需求对书本进行了分类管理，有关太阳能电池的书在这个书架上，而与法国文学有关的书则在那个书架上等。而对于一个读者没有偏好，只能随机游走的图书馆来说，只要走错一步就会灰飞烟灭，那么谁都不敢在这样的图书馆里随便走动，读者只能停留在眼前的书架上。如此一来，它们就成了鼠目寸光的伪学者，除了精通自己所在的书架之外，对其他领域一无所知，也不会学到任何新的东西。这可不是在这个

多变的世界上生存下去的好办法。对于这样的读者来说，代谢图书馆简直是专门为它们寻求新性状设计的。

更奇妙的是，其他与生命有关的自然图书馆也遵循相同的组织方式。

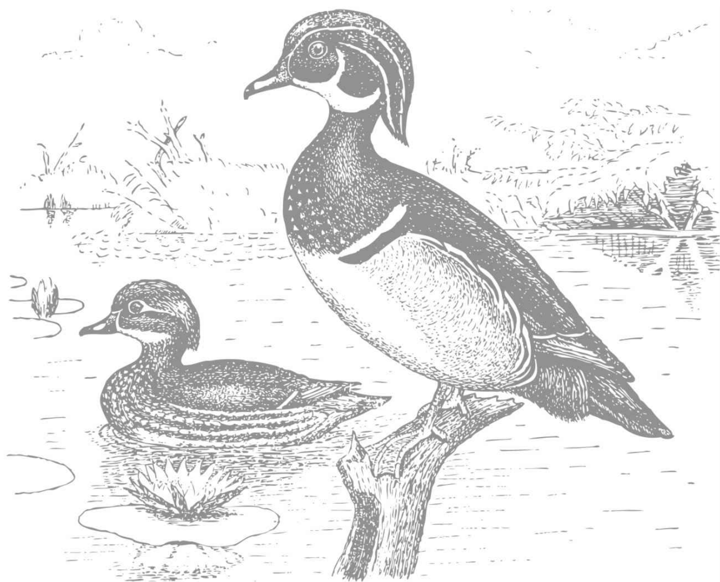


ARRIVAL OF THE FITTEST

04 构型之美

Solving Evolution's Greatest Puzzle

蛋白质是生命的驱动者。每种蛋白质的构型都高度复杂，与它们执行的功能相适应。蛋白质的构型维持着生命世界的运转。大自然可以用蛋白质书写不同的文本，更多的文本就意味着更多的构型，参与更多的催化反应，执行更多的功能和完成更多的任务。



北极鳕鱼体型细长，体表呈褐色，腹部为银白色，鳍则是黑色，长约18～30厘米，是一种毫不起眼的海洋生物，但有一点除外：北极鳕鱼生存繁衍的区域位于海平面以下900米，纬度和北极相差不超过6度，那里的水域常年水温在0摄氏度以下。

在这样的温度下，大多数生物体内的液体会变成冰晶，边缘像锻造精良的利剑一样美丽，也像利剑一样致命，因为冰晶切割活体组织就像切黄油一样毫不费力。恒温动物体内具有温度调节系统，因此即使身处零度以下也能生存。但鱼类没有调节体温的能力，即便如此，北极鳕鱼在零度以下依旧能够生存。

北极鳕鱼能够生存的秘诀是体内合成的抗冻蛋白。抗冻蛋白降低了体液的冻结温度，很像汽车引擎冷却剂里的防冻剂。功能繁多的不同蛋白质是自然进化能力的典型例子。只要通过改变氨基酸序列以产生特定的蛋白质，地球海洋中的大片生命不能涉足的区域就可以变成适宜生存的乐园。

抗冻蛋白只是无数种进化奇迹之一，类似的奇迹在鱼类和所有其他生物的细胞中普遍存在。如果你能变小，在细胞里穿梭，你肯定会惊讶于细胞里居然有这么多不同种类的分子，数以百万计。有小些的，如水分子，有的稍大一点，譬如糖或者氨基酸，也有更大的大分子，如蛋白质。所有这些分子推推搡搡，你挤我我挤你，就像高峰时刻挤地铁上下班的乘客。

蛋白质是细胞成分分子中的庞然大物，是生命的驱动者。我们已经认识过代谢酶了，它们通过对相对较小的分子进行聚合、剪切或仅仅进行原子重排，合成细胞所需的所有物质，包括这些代谢酶自身的氨基酸。不过，并不是所有的蛋白质都是酶。有些蛋白质是细胞内的动力来源，例如驱动蛋白，驱动蛋白沿着细胞间纵横交错的刚性分子“电缆”移动，运输含有各种分子物质的小囊泡。一旦这些在细胞内“跑长途”的运输分子停止工作，混乱和无序就会接踵而至。举个例子，有一种驱动蛋白负责运输构建神经细胞之间连接的物料分子，倘若它的基因发生变异，会导致一种叫2A型腓骨肌萎缩症（Type 2A Charcot-Marie-Tooth disease）的绝症，表现为手脚无力并伴有感觉障碍。

还有一些蛋白质附着在DNA上，是基因的开关，这些调节性的蛋白质负责把基因编码的信息转化成氨基酸链。上百个类似的调节蛋白往往

会同时展开工作，每个调节蛋白总是对某些基因产生作用，而对另一些则没有影响。（调节蛋白本身也是生物进化的动力之一，我们将在第5章讨论。）

蛋白质的功能还不止于此，细胞内还有组成细胞骨架的支柱蛋白质、输入养分的蛋白质、将废物排出细胞的蛋白质以及在细胞间传递分子信息的蛋白质等。

每种蛋白质对于生物表现型的塑造都有它独特的功能，而对于蛋白质来说，构型是它们最重要的特征。这里所说的构型不仅指蛋白质中20种氨基酸本身的分子形状，以及氨基酸之间相互的连接方式——这些统称为蛋白质的一级结构，还代表线性的氨基酸链经过空间折叠形成的立体结构。蛋白质的空间折叠我们在第1章中就已经略有涉及。

亲水性氨基酸倾向于靠近它们周围的水分子，而疏水性氨基酸——就像生物膜成分中的脂质，则倾向于逃避周围的水分子，氨基酸分子对水的不同亲和力使蛋白质的一级结构能够以特定的方式折叠。在热力学震动的驱动下，氨基酸链在折叠过程中会尽可能地尝试所有可能的构型，最后，大多数疏水性氨基酸会聚合在一起，形成紧密的核心。而最外层则被亲水性分子所覆盖，包裹内部的疏水性核心。

另外，有些氨基酸能够互相吸引，有些则会互相排斥，这些化学作用力也影响了蛋白质的折叠方式。蛋白质折叠过程仅由蹦蹦跳跳、行踪不定的分子决定，这再次提醒我们，自组织能力于生命而言有多么重要。我们体内有上万亿个细胞，只要细胞内形成一条新的蛋白链，它都要发生空间折叠。所以在一天中，蛋白质的折叠在每个细胞内都要上演几百万次。

如果从原子层面看，发生空间折叠的蛋白质看起来就像一团没有形状的泡沫，比如第2章中提到的能水解双糖的蔗糖酶。我们不妨退一步，把注意力集中在氨基酸链上（见图4-1），就可以辨认出些许有规律的氨基酸空间排列模式，这种模式在许多蛋白质中都存在，包括了一种螺旋形的结构和一种扁平的片状结构，片状排列也叫β-折叠。α-螺旋和β-折叠是蛋白质折叠的主要方式，两者即蛋白质的二级结构。多个α-螺旋和β-折叠，连同联结两者的非二级结构部分，构成了图4-1中所示的蛋白质的复杂三维结构，我们把这种三维结构称为蛋白质的三级结构。

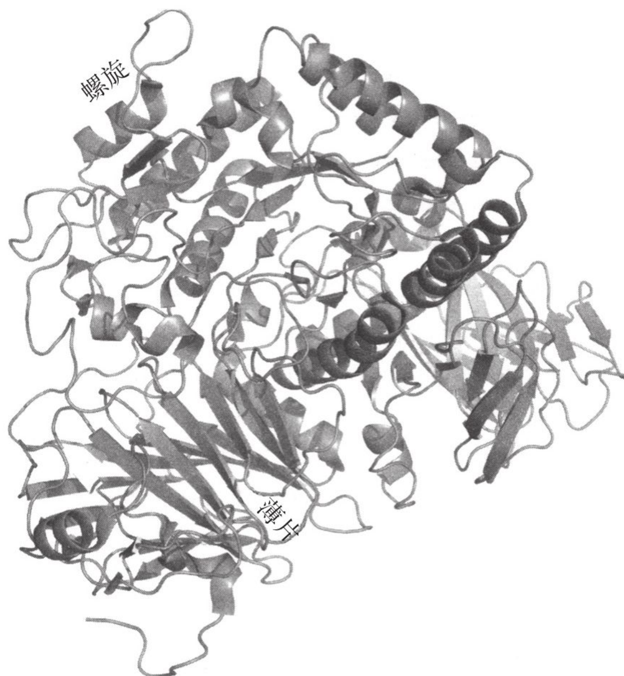


图4-1　完成空间折叠后的蔗糖酶

虽然图4-1中的折叠看起来很像一团乱糟糟的意大利面，不过这种折叠方式实际上是高度有序的：任何一个蔗糖酶中的两条氨基酸链折叠的方式都是自发完成的，且折叠的产物完全一致。构型对于保证蛋白质的功能至关重要：热能导致折叠的蛋白质分子不断振动和振荡，而α-螺旋和β-折叠则起到引导并限制分子热运动的作用。振动受限让蔗糖酶这样的酶能够催化糖的裂解反应，原理有点像剪刀：如果没有连接刀片的转轴限制它们的运动，剪刀也就无法裁纸。鉴于热运动对酶分子的重要性，所以对每种酶分子的催化作用而言，都存在一个最适的理想温度：热量太低，分子振动微弱，不足以组织分子运动；热量太高，剧烈振动则会使空间折叠分崩离析，导致蛋白质变回线型氨基酸链。更糟的是，未折叠的蛋白质经常聚合成大团大团的惰性物质，就像熟鸡蛋里的蛋白。未折叠的成团蛋白质不仅无用，而且有害。就像如果你的大脑里积累了太多蛋白质块，就会引起严重疾病，例如阿尔茨海默氏症。

蔗糖酶和其他蛋白质在振荡中形成的构型复杂多样，且各自都有着特定的功能。每种蛋白质的构型都高度复杂，与它们所执行的功能相适应。用达尔文描述生命世界的话来说，这是一个“无尽之形最美”（**endless forms most beautiful**）的世界。蛋白质的构型维持着生

命世界的运转。

蛋白质不仅需要处理眼下的工作和任务。如同人类的经济社会一样，生物也需要面对瞬息万变的挑战。作为应对，进化为生命带来了新的蛋白质构型，而具有新构型的蛋白质则可以承担新的工作。每当生命需要解决新问题时，比如在极度低温的环境中，体内生长的冰晶变成致命的刀片，威胁到自身的生存时，新的招募工作就开始了。

无论是早先的高炉，还是如今的智能手机，人类社会中的发明往往需要经历漫长的独立研发过程，并非一蹴而就。与之类似，自然界塑造生物新性状的过程也往往不是瞬间实现的。抗冻蛋白就是一个例子，不只是北极鳕鱼体内有抗冻蛋白，南极鱼类也有，但是两者的抗冻蛋白分别起源于不同祖先体内的两种蛋白质。抗冻蛋白甚至有过不止一次起源，不仅如此，有些鱼类还进化出了不止一种抗冻蛋白。美洲拟鲽是北大西洋的一种比目鱼，它的体内能够合成两种抗冻蛋白，一种防止血液结冰，另一种防止皮肤结冰。以进化的角度来衡量，有些蛋白质的出现着实非常迅速，只用了不到300万年的时间。

某些早期生物对冻伤非常敏感，它们体内的蛋白质与抗冻蛋白的氨基酸序列往往相去甚远，不过蛋白质进化所需的变化通常比我们想象的要少得多。只要改变一个氨基酸，合成组胺酸的酶就会变成功能不同的另一种酶，合成的产物也变成了色氨酸。大肠杆菌的一种酶能催化并从阿拉伯糖[16]中摄取能量，改变这种酶中的一个特定的氨基酸，它的功能就从转移酶变成了裂解酶。

可见，微小的变化同样能够对生命造成巨大影响，中亚的斑头雁也是一个例证。斑头雁是世界上飞得最高的鸟类之一。它必须飞得足够高，因为它的迁徙路线需要途经珠穆朗玛峰，那里的海拔超过8 000米。在这个高度上，周围的空气非常稀薄，鸟儿必须更加拼命地拍打翅膀，而且那里的氧气含量仅为海平面的1/3。攀登珠穆朗玛峰的登山者在到达这一海拔时往往需要借助氧气罐，而乘坐喷气式飞机的乘客则需要加压舱的保护。大雁无法借助这两者中的任何一项技术，但是没关系，它有更好的办法。斑头雁体内的血红蛋白在氨基酸序列上发生了变异，这种蛋白质负责将氧气从肺输送到肌肉。与我们体内的血红蛋白相比，斑头雁体内的血红蛋白和氧气结合更紧密。斑头雁能从稀薄的空气中摄入氧气分子，在别的鸟类不得不因为过高的海拔着陆时，斑头雁却能继续飞行。

北极鳕鱼体内的抗冻蛋白和斑头雁体内的氧合血红蛋白，都是弥足

珍贵的进化产物。它们拓展了生物的活动范围，而更大的栖息地则意味着更多的食物、更高的生存概率以及更多的后代。还有一些进化的优势成果与这些改良的分子略有差异，比如区别一种食物与另一种食物的能力，比如晚餐选择有营养的而不是有毒植物的能力。这些性状优化的是生物的知觉，而不是它们的机动性。人类眼球后方的视网膜里含有三种视蛋白，这三种视蛋白在感光和适应不同波长的能力上高度特异化。多亏有它们，我们才能看见色彩斑斓的世界。但这一切并不是从生命出现之初就与生俱来的。

我们最古老的脊椎动物先祖可能只拥有一种视蛋白，它们眼里的世界是黑白的。大多数哺乳动物有两种不同的视蛋白：视红蛋白和视蓝蛋白，它们能看到的世界主要基于这两种不同的颜色。但我们和黑猩猩等近亲则能够看到基于三种原色的世界，也许这是因为色觉有助于我们的祖先觅食：在绿叶的衬托下，水果往往更加显眼。当然，不管什么原因，色觉进化所需的变化很少，少到只要改变三个氨基酸就能把视红蛋白变成视绿蛋白。

色觉进化等改变对我们往往是有益的，但也有一些进化对我们是有害的，例如某些细菌进化出的对医生开出的抗生素产生耐药性的能力。尽管我们不断地改进着抗生素，但是耐药性却始终不可避免，这是一场发生在细菌与生物技术专家之间的军备竞赛。这场竞赛让人想起刘易斯·卡罗尔（Lewis Carroll）在《爱丽丝镜中奇遇记》（*Through the Looking-Glass, and What Alice Found There*）中描写的红皇后，她曾对爱丽丝说过这样一句广为人知的话：“你瞧，在我们这儿得拼命地跑，才能保持在原地。”细菌在这场竞赛中找到了各种各样的新蛋白，有些蛋白质能够破坏抗生素分子，另一些则被称为“外排泵”，它们能把抗生素强行排出细胞，就像一支细菌救援小队把有毒气体排出被污染的房子。

基因转移加上频繁的人口流动可以在几个月之内把这些优良的性状带到世界的每个角落。还有一些蛋白质尤为阴险，它们负责排出的抗生素不止一种，如此一来，细菌就同时对多种抗生素产生了耐药性。说来奇怪，如果我们自己身体里的细胞失控，疯狂繁殖时也常常利用相似的外排泵对抗它们厌恶的抗癌药物。这不仅仅是癌细胞和细菌在面对危险时的英雄所见略同，同时它也是我们在对抗癌症的战斗中屡战屡败的原因。

生物进化中出现的外排泵并非无中生有，而是对已有运输蛋白进行修饰的产物。运输蛋白对维持细胞的日常运作与生存至关重要，因为它

们不停地将成千上万的分子（养分、废物、建筑材料）送往细胞内的不同目的地。那么我们真的应该称之为新蛋白吗？同样的疑问也存在于斑头雁体内的改良血红蛋白和灵长类体内的感光细胞中。大自然不过只是胡乱摆弄了几下原有的血红蛋白，让它和氧气结合得更紧密；或者胡乱修改了一下视蛋白，调整了它的色觉灵敏谱而已，两者都算不上是严格意义上的“新”蛋白。

不过，如果你试想一下这些变化带来的影响：一只可以穿越任何山脉的鸟迎来了几百万平方公里的新栖息地；我们看到的世界是黑白的该有多无趣；耐药或者不耐药对细菌而言意味着生死之别。单就差异巨大的结果而言，这些小小的变化就足以被称为新性状。上述的例子无不说明，只要稍微改变几个原子就可以影响比原子大几百万倍的生物，并永远改变这个生物后代的命运。



我们在第3章里看到，通过基因水平转移重组代谢酶，大自然不断更新着生物体内的生化反应。即便如此，代谢酶自身也并不是这样形成的。从上述最后举的几个例子中我们可以看出，大自然会通过改变已有蛋白质的氨基酸序列创造新的蛋白质，我们已知的5 000多种酶中的每一种都是按照这种方式被创造出来的。此外，还有不胜枚举的蛋白质在我们体内负责调节基因、输送物质、收缩肌肉、运输氧气、输入养分、排泄废物、在细胞间传递信息以及承担其他无数种任务。蛋白质多样的功能足够我们写上整整一本书。事实上已经有这样的书了，这种专业的书对某些蛋白质的记述非常完备，事无巨细。

但这本书不属此类。

你无法只是通过道听途说就了解有一种蛋白质叫抗冻蛋白，还有一种叫视蛋白，以及理解这些蛋白质起源的真相，就像你没法只靠几个国家的卫星图像就画出一张完整的美国地图。要解释新蛋白质的起源需要我们拿它们与大量原始蛋白质进行比较，成百上千对地进行比较。

如果能解读基因的DNA或者基因编码的氨基酸链，也就是蛋白质的基因型，这个任务就会容易一些。对两者进行解读的先驱之一是英国生物化学家弗雷德里克·桑格（Frederick Sanger）。他是极少数得过两次诺贝尔奖的科学家之一，第一次是因为他成功破译了胰岛素的氨基酸序列，而第二次则是表彰他成功完成了对DNA的碱基测序。在距离他做出

卓越贡献的几十年后，我们才终于掌握了解读代谢基因型的技术，才认识了更多蛋白质的基因型和表现型。这些基因型和表现型来自各种生物，所处环境各异：北极圈荒原、热带丛林、山麓、深海、人体内脏、滚烫的温泉、贫瘠的荒漠、肥沃的平原、肮脏的水沟和清澈的河渠。

这么大一堆关于蛋白质的事实如果不经组织，简直就是一本疯子编纂的字典，里面的几百万个单词杂乱无章，毫无头绪可言。然而一旦经过组织，这些事实就成了图书馆的一部分，这个图书馆和第3章中巨大的代谢图书馆类似。这个宇宙图书馆里收录的正是蛋白质的基因型，每个文本都由20个字母构成的字母表写就，每个字母对应一个氨基酸。这座图书馆收集了生命已经创造和能够创造的所有蛋白质，有时也被称为蛋白质空间（protein space）或序列空间（sequence space）——因为每个文本都对应一个唯一的氨基酸序列。

通过与前面章节类似的计算方式，我们大概已经可以预见，这个图书馆的规模和代谢图书馆一样惊人。回想一下，20种可能的氨基酸，两个字母可以构成的文本就有400（20^2）种。同理，3个氨基酸构成的可能文本有8 000（20^3）种，4个氨基酸对应于16万（20^4）种文本，以此类推。像这样的短文本充其量只能叫肽，大多数蛋白质包含的文本要长得多，即多肽，它的可能文本数量随着长度增加呈爆炸式增长，即便是仅由100个氨基酸构成的蛋白质，可能的文本数量也已经超过了10^{130}种。这个数字大得难以想象，但是蛋白质图书馆内的馆藏数量比这还大，因为像蔗糖酶这样的蛋白质含有1 000多个氨基酸，而有些人类蛋白质比蔗糖酶还要再长许多倍。其中有一个庞然大物叫肌联蛋白（titin），含3万个氨基酸，是肌肉的弹性成分。由此可见，蛋白质图书馆的规模同样是超宇宙常数级别的。

蛋白质图书馆和代谢图书馆的相似之处不仅在于规模。和后者一样，蛋白质图书馆也是一个超几何体，相似的文本彼此邻近。每个蛋白质文本位于这个超立方体的一个顶点，就像在代谢图书馆里一样，每个蛋白质都有许多直接相邻的邻居，这些邻居和它只差一个字母，位于超立方体上相邻的顶点。

我们以一个含有100个氨基酸的蛋白质分子为例，如果你想改变蛋白质的第一个氨基酸，那么就有19个选项，这也就意味着与这个蛋白质只差第一个氨基酸的蛋白质邻居有19个。按照同样的思路，这个蛋白质有19个与其第2个氨基酸不同的邻居，19个与其第3个氨基酸不同的邻居，19个与第4个……19个与其第100个氨基酸不同的邻居。一句话，

我们的蛋白质有1 900个直接邻居。这样一个社区已经很庞大了，但如果你改变的不是一个氨基酸，而是两个或更多，那么这个社区还会更大。显然，对进化来说这不是坏事：只要简单改变一个或几个氨基酸，就可以产生许多新的蛋白质。

在这座图书馆迷宫里漫游，要是手里没有一团展开的毛线丈量走过的路程，很容易就会迷路，这一点也和代谢图书馆类似。在这里我们也需要借助某种方式来衡量蛋白质图书馆里的“距离”，于是我们采用了两个蛋白质相异的氨基酸数目作为衡量距离的单位。这个标准可以告诉你，从一个蛋白质文本到任一其他文本要走多远，即需要改变多少个氨基酸。

图书馆中的文本很重要，但更重要的是每个文本承载的意义。我们的双眼无法解读这种意义，无法阅读蛋白质化学语言的单词、句子和段落，但生命自身精通这门语言，并能分辨出一个蛋白质文本到底是文风优美的佳作，还是词不达意的垃圾。

细胞判断蛋白质是否有意义的标准很实际：能让细胞存活的蛋白质就有意义。只有有用的蛋白质才有意义，有缺陷的变异蛋白不能正确完成折叠，自然也就一无是处。如果“意义”这个词听起来过于以人类为中心，我们不妨参考一下符号学——一门语言学的分支，主要研究意义的意义，其中对“意义”的定义是任何符号（随便什么东西，可以是路标，可以是一本书）所指涉的内容。根据这个定义，如果蛋白质的基因是符号，那么它所编码的蛋白质氨基酸序列以及蛋白质在细胞内所起的作用就是它的意义。

宇宙图书馆里到底确切地藏有多少本有意义的书，我们仍然无从得知。但经过几十年的研究，如今我们已经可以估算蛋白质图书馆里有意义的蛋白质的数量，因为大多数有用的蛋白质都有特定的折叠形状。从图书馆里随机选取一个书架，随机选取一个蛋白质，它能够折叠的概率至少是万分之一。这个概率听起来好像不是很大，但请记得，宇宙图书馆本身非常巨大，光是由100个氨基酸组成的蛋白质就超过10^{130}个。即使其中只有万分之一的蛋白质能够折叠，也有10^{126}个，即1后面跟着126个0，这比全宇宙中的氢原子数量还多。由此可知，有意义的蛋白质的数目大得超乎想象。



进化会利用大量的生物体对蛋白质图书馆进行探索。DNA一代接一代地复制，难免会出现复制错误，改变DNA链上的碱基，如腺嘌呤变成胞嘧啶，胸腺嘧啶变成鸟嘌呤，或者发生其他变化，每改变一个氨基酸，蛋白质就会发生改变。变化后的文本可能具有全新的用途，想要理解这一过程，我们就得绘制蛋白质图书馆的地图，就像在代谢图书馆中做过的那样。这个任务没有看上去那么难：多亏研究蛋白质的科学家们在过去数十年中的不懈努力，我们已经知道了成千上万种蛋白质的折叠方式、功能以及在图书馆中的位置。另外，借助20世纪的分子生物学技术，我们可以从书架上取下任意一卷书，合成相应的蛋白质，并在实验室里研究它的折叠方式和功能。

有关蛋白质进化中的一个最简单的问题，我们在之前的章节中已经探讨过了。要找到一个有一丁点意义、有助于生物体存活的蛋白质有多难呢？如果图书馆里只有一个这样的蛋白质，即便从遥远的大爆炸开始找也很难找到。既然存在大量有意义的蛋白质，那么生命在面对不同的挑战时，就会有多种解决方式，但到底有多少种呢？

2001年，哈佛大学的安东尼·基夫（Anthony Keefe）和杰克·绍斯塔克（Jack Szostak）试图回答这个问题，他们研究的蛋白质家族的重要性不亚于生命历史中出现的任何其他性状：这类蛋白质可以联结三磷酸腺苷，而我们已经在第2章中说过，三磷酸腺苷是生命的电池。一般情况下，蛋白质通过裂解三磷酸腺苷摄取工作时所需的能量，包括运输材料、收缩肌肉、构建新分子等。

想要释放和利用三磷酸腺苷的能量，蛋白质首先要结合三磷酸腺苷。如果庞大的蛋白质图书馆里只有一种蛋白质能够结合三磷酸腺苷，盲目寻找不过是在白费力气，想找到它除非有奇迹出现。基夫和绍斯塔克想弄清楚，图书馆里能够结合三磷酸腺苷的蛋白质究竟有多稀有。他们采用化学手段创造出许多不同种类的蛋白质，每种蛋白的氨基酸序列都不相同，完全随机。这个人工设计的过程相当于从蛋白质图书馆书架上随机取下一册书卷。研究者制造的随机蛋白质均含有80个氨基酸，这样的蛋白质数量超过10^{104}种，不可能全部在实验中合成出来，但这个实验中所合成的随机蛋白质数量已经相当惊人了：大约有6万亿种。

基夫和绍斯塔克发现，其中4种毫无关联的蛋白质可以结合三磷酸腺苷。6万亿种蛋白质中有4种可以结合三磷酸腺苷的蛋白质，是不是听上去一点都不富裕？但是按照这一比例，所有包含80个氨基酸的候选蛋白质中能结合三磷酸腺苷的就多了去了：10^{93}个。三磷酸腺苷结合问题

的答案是一个天文数字。

麻省理工大学的约翰·里德哈尔-奥尔森（John Reidhaar-Olson）和罗伯特·索尔（Robert Sauer）研究了同样的问题，但采用的方式不同。他们研究的是一种调节蛋白，这种蛋白可以关闭病毒的基因，后者能感染细菌。这种病毒名为A字形噬菌体，它的DNA所编码的蛋白质使它能自我复制并杀死宿主。但利用开关关掉基因后，这种病毒也可以在细菌体内休眠，直到时机成熟再自我复制并杀死宿主。通常当宿主遭遇不幸时，如饥饿、抗生素污染、过量的紫外线照射，就意味着时机到了，病毒于是趁机开始复制，复制所得的子代病毒冲破细胞，将满目疮痍的细菌遗弃。那场面用一句不恰当的俗话说，大概就是“树倒猢狲散”。

奥尔森和索尔探索了蛋白质图书馆中这个病毒开关附近的一个社区，随机创造出大量社区中的氨基酸序列，从而探究哪种序列可以产生有效的开关关闭病毒基因。他们计算的结果是，整个图书馆中有超过10^{50}个文本能够编码关闭基因的分子开关。他们把类似的算法应用于另一种蛋白质，即合成氨基酸所需的酶，发现了多达10^{96}种蛋白质可以完成这项工作。

自然界中的抗冻蛋白给了我们一个提示，之后我们在实验室中则以实验证实了这个猜测：无论是结合三磷酸腺苷，关闭病毒基因，还是催化生化反应的蛋白质，都不是唯一的，应对相同问题的解决方案甚至可能超过100万种。具有不同功能的蛋白质数目是一个天文数字，每一种都对应蛋白质图书馆中的一卷书。图书馆中馆藏数量之多，难以想象。就生物的创造力而言，只有我们想不到，没有自然界做不到。

事实上，解决特定问题的书在图书馆里取之不尽。当然，知道这一点还不够。我们还要找出这些答案的位置和组织方式，它们是整整齐齐地排在书架上呢，还是随意摞成一堆？仅有实验室的实验是远远不够的，因为即使实验能够合成并测试惊人数量的蛋白质，但与自然界中实际的蛋白质数量相比，依旧显得无足轻重。在自然界中，每一天都有不计其数的生物体在加班加点合成新的蛋白质，每个生物体都是合成蛋白质的量产工厂，而每一个蛋白质都不过是在持续了亿万年的蛋白进化之路上，最后的那一个脚步而已。

蛋白科学家早就已经注意到了蛋白质的多样性。如果把蛋白质比作糖，拥有数量庞大的蛋白质的自然界就像一家巨大的糖果店，心怀热忱的科学家就像孩子一样一拥而上。比起实验室中得来的数据，科学家在成千上万的生物体中得到的有关蛋白质进化的知识要多得多。我们前面

探讨过的斑头雁体内负责输送氧气的血红蛋白就是一个很好的例子。

血红蛋白的功能不难理解，它往返肺与身体组织之间，完成对氧气的结合或释放，它的重要性也无须赘述。血红蛋白属于一个结合氧气的蛋白质家族，即球蛋白（globins）。球蛋白不仅对我们，对于许多其他的哺乳动物、鸟类、爬行类和鱼类也同样重要。这些物种最初具有共同的祖先，但随后经过无数代遗传，如父辈、子辈、孙辈和数不清的曾孙辈，在一代代演替过程中，编码血红蛋白和所有其他蛋白质的DNA经历了无数次复制。虽然每次复制都极少出现复制错误，对我们的细胞而言，DNA复制过程中大约平均每4 000万个碱基中才会出现一个复制错误，但只要假以时日，只要时间足够长，理论上一个基因组中的所有基因都会出现复制错误，从而导致它们编码的蛋白质产生变化。

编码氨基酸序列的基因变化后，球蛋白不再按照正确的方式折叠，氧气也就不能流向需要的地方。简单地说，这通常意味着死亡。但发生改变的蛋白质不一定会完全丧失原本的功能和意义。甚至有些改变既不会改变蛋白质的功能，也不损害基因存在的意义，并能够传给下一代。经历无数代繁衍，复制错误，尤其是某些可容忍的复制错误在基因组中逐渐积累，就会慢慢改变蛋白质的氨基酸序列。

图4-2展示的是人类以及3种物种分类上的亲缘生物，各自血红蛋白中的10个氨基酸片段。图中的每个字母都取自包含20个字母的字母表。科学家通常用字母缩写代表氨基酸：V代表缬氨酸，A代表丙氨酸，诸如此类。大约500万年前，即差不多20万代人之前，我们和与我们血缘关系最近的黑猩猩还拥有一个共同的祖先。由于这段时间还不足以积累大量的遗传错误，因此黑猩猩的球蛋白文本迄今没有改变太多。在图4-2中展示的球蛋白片段里，人类和黑猩猩只有一处不同：在人类血红蛋白为丙氨酸（A）的位置上，黑猩猩的则为谷氨酸（E）。

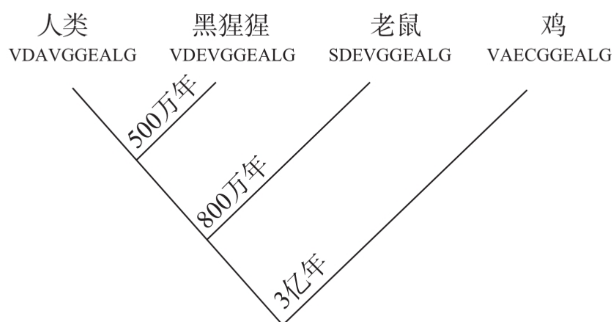


图4-2　蛋白随时间变化图

大约800万年前，人类的祖先和老鼠的祖先分道扬镳。因此比起黑猩猩，老鼠的球蛋白累积变化的时间更长。图4-2显示，人类和老鼠有两处氨基酸差异。鸡的祖先和我们的祖先分开得更早，是在近3亿年前，所以其间累计的氨基酸差异则相应达到了6个。

还有几百万种生物体含有血红蛋白，除了恒温的脊椎动物，还有爬行类、青蛙、鱼类、海星、软体动物、苍蝇、蠕虫，甚至植物。其中一些物种从生命之树的同一枝条上生长出来，距离它们拥有同一个祖先的日子还不太远。它们的球蛋白基因在生物史上的大部分时间里都是相同的，只是在最近才分道扬镳，但是依旧十分相似。另一些生物体位于生命之树的不同树枝上，距离它们拥有共同祖先的日子较久远，控制球蛋白合成的基因相差也更大。但不管这种差异有多大，它们编码的球蛋白都能正常工作，否则这些球蛋白基因就不会存留至今。每个幸存的基因解决氧气结合问题的方案都不尽相同。生命每延续1　000年，就会进入蛋白质图书馆的更深处，在随机进化之旅中探索全新的球蛋白文本。

想知道球蛋白在进化之旅中走了多远，需要想想我们最远的亲戚：植物。尽管植物没有血液，但事实上还是有一些植物能够合成球蛋白。

大豆、豌豆、苜蓿等豆科植物可以从空气中吸收至关重要的氮元素，而空气中的氮几乎是取之不尽的。（其他植物大多需要从土壤中吸收氮，除非农民施用化肥，否则土壤的氮含量通常很低。）豆科植物借助细菌从空气中吸收氮，这种细菌成群结队生活在植物根部附近，体内含有一种特殊的酶，能将空气中的氮气转化为铵盐，铵盐也是含氮化肥的主要成分。这种天才的共生关系只有一个缺点：大气中的氧气会破坏固定氮气的酶。植物为保护这种酶而合成球蛋白，根部的共生细菌因而得以远离氧气。

植物和动物的共同祖先可以追溯到10亿多年前，动植物处于生命之树不同的主枝上，两者间的球蛋白的差异大得惊人，这表明动物和植物相互独立的进化旅程已进行了很久。举个例子，羽扇豆和昆虫的球蛋白中几乎有90%的氨基酸都不同。然而，如图4-3所示，这些球蛋白不仅都能结合氧气，折叠形状也十分相似。左图是一种豆科植物的球蛋白折叠，右图是一种小型双翅昆虫摇蚊的球蛋白折叠。两种蛋白质都有几处螺旋体一样的结构，例如左上方和右下方有两处排列十分相似的平行螺旋。图像无法完全说明这两个球蛋白有多相似，如果你转动分子，把其中一个覆盖到另一个分子之上，就会发现两个分子中原子的空间分布几乎一模一样。尽管已经各自独立进化了10亿多年，但这些球蛋白的折叠

方式依然十分相似。

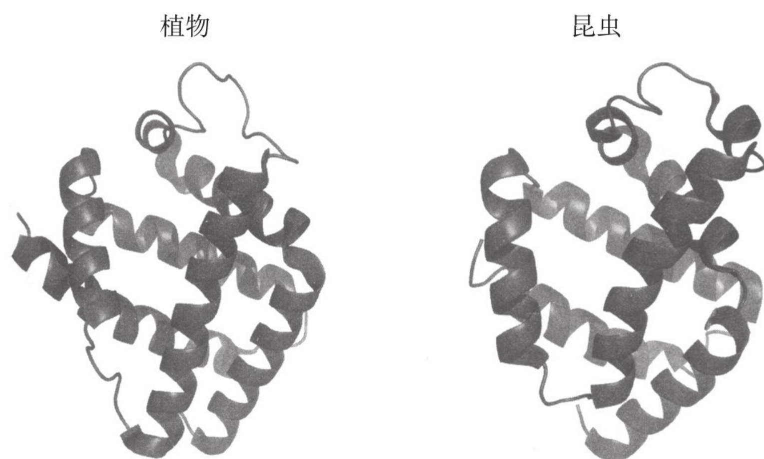


图4-3 两种折叠方式相似的血红蛋白

豆科植物与昆虫球蛋白的氨基酸差异极大，但这种差异稀松平常。哪怕对于两种动物球蛋白而言，很可能有80%的氨基酸也是不同的，比如蛤和鲸。尽管有种种差异，上述以及其他生物体体内的上千种球蛋白仍然彼此相关联，组成一张网络，遍布蛋白质图书馆。这张网络上的通路连续不断，从共同祖先出发，每走一步改变一个氨基酸，但文本含义保持不变。

类似的基因型网络我们已经在代谢图书馆中探讨过了，无论生物进化在这张网络中往哪个方向走，走多远，代谢表现型的意义总是保持不变。进化在探索蛋白质图书馆的过程中采用了一种不同的策略，不是基因水平转移，而是改变单个氨基酸，但两者的本质是相同的。基因型网络将不同的球蛋白连接在一起，网络的根须触手一直延伸至蛋白质图书馆的深处。进化可以沿着这个网络探索图书馆，不致迷路而陷进由无用分子围成的致命流沙中。

球蛋白形成的基因型网络规模庞大、交错纵深并不是例外，而是普遍规律。折叠方式、催化反应以及祖先相同的酶，它们的氨基酸的相似度通常不超过20%。我们能认识到这一点，是因为科学家已经在图书馆里确定了上千种已知酶的编码文本位置。通过给文本编目，我们能够绘出酶的基因型网络在图书馆中的通路，其中有一些甚至比球蛋白延伸得更远。延伸最远的是TIM桶状蛋白质，TIM是磷酸丙糖异构酶（triose phosphate isomerase）的首字母缩写。由于α-螺旋和β-折叠的排列方式

像木桶夹板，所以这种酶被称作桶状蛋白。TIM有助于从葡萄糖中摄取能量。令人震惊的是，有些以相同方式进行折叠的酶与TIM没有一个氨基酸相同。它们分别位于蛋白质图书馆的对角位置，即所有字母都互不相同，尽管如此，它们却携带着相同的化学信息。这些蛋白质就像无数个不同版本的《哈姆雷特》，尽管不同版本的4 000行台词中只有几百行相同，甚至没有一行相同，但每个版本都完整地讲述了莎翁笔下那个王子复仇的故事。

自然实验室中的上千种蛋白质同样叙述了一个类似的故事：不管是酶、调节因子还是像血红蛋白那样的运输分子，当我们需要一个新的蛋白质解决眼前的问题时，解决方案往往多得数不过来。不仅如此，应对相同问题的蛋白质由一张众多蛋白质文本构成的巨网相连，遍布蛋白质图书馆。在某些蛋白质网络中我们已经能够认出数千种蛋白质了，可是这也只是沧海一粟，要知道，一张网络中具有相同表现型的蛋白质往往多达数万亿个。

有些未知的蛋白质属于早已灭绝的生物，但是绝大多数蛋白质甚至从未在自然界出现过。生命历经的40亿年太短，只够创造出10^{50}种蛋白质，这只占蛋白质图书馆所有文本中的极小一部分。不论巨大的生命之树上挂着多少蛋白质，也不论这棵树有多么高大、多么美丽，它终究只是脏兮兮的镜子里污迹斑斑的影像，是柏拉图的理想世界中模糊不清的幻影，唯有背后那张更大的基因型网络才是这一切的本质。



我们在第3章里看到，在进化过程中，有几十亿读者通过基因型网络探索着代谢图书馆各个角落里的不同社区。尽管有些探索者掉下网络一命呜呼，但也有一些探索者通过网络发现了新表现型的进化文本。基因型网络或许同样可以服务于蛋白质，前提是蛋白质图书馆里的社区也具有多样性。否则，进化的蛋白质还不如待在原地不动。因为如果图书馆不同区域堆放的书籍相同，也就没有探索图书馆的必要了。

图书馆中每个蛋白质附近书架上的文本是否意义相近，是否就像现代社区中大同小异的家家户户？还是更像中世纪的村庄，风格独特，魅力各异，所含蛋白质拥有独一无二的新功能？尽管我们研究了几十年蛋白质，如今甚至可以用计算机挖掘堆积如山的蛋白质数据，但是对于这个问题，直到现在我们也没有找到答案。

要回答这个问题，光有计算机还不够，还需要热爱书本的图书管理员。一个年轻的智利研究者埃万德罗·费拉达（Evandro　Ferrada）带着这份热爱来到苏黎世，加入了我们的研究小组攻读他的博士学位。埃万德罗曾经有过研究蛋白质的经验，他能熟练地从巨大的蛋白质数据库中筛选出所需的信息，不论是蛋白质的折叠方式还是蛋白质分子内的原子信息。埃万德罗个性安静，常常陷入沉思，这种个性我以前在别人身上见过，这些人常常与生命的深层奥秘缠斗。也许这也是他同意研究这个问题的原因，因为蛋白质的空间结构正是这样的奥秘：挑战性强，意义深远，异常艰难，但也不是没有被解决的可能。另外，蛋白质的空间结构里还隐藏着蛋白质进化的秘密。

埃万德罗重点研究的是酶，因为这类蛋白质的种类异常丰富。这毫不意外，因为酶需要催化5　000多种不同的化学反应。科学家已经对酶进行过深入研究：他们已经把蛋白质图书馆中散落的几千种蛋白质标记了出来。一旦我们知道了酶的精确位置，就可以用计算机进行分析。埃万德罗利用电脑选出一对蛋白质，它们的折叠方式相同，但在基因型网络中处于不同位置。紧接着，他检验了第一个蛋白质所在社区的一小部分相邻蛋白，列出了其中所有已知的蛋白质和它们的功能。之后，他以相同的方式检验了第二个蛋白质所在的社区，列出了经过检验的所有已知蛋白质和它们的功能。他比较了两张列表，关注点集中在两个列表中的蛋白质是否相同，以及两个社区的蛋白质功能是否相同。然后他选取了另一对蛋白质，重复相同的检验过程，关注相同的问题，直到他研究了几百对蛋白质和它们所在的社区。

最终的答案很简单：即使两个蛋白质在图书馆里离得很近，它们的社区内包含的大部分蛋白质功能也不同。比如，某两个蛋白质中存在差异的氨基酸只有不到20%，即便如此，它们各自所在社区里的蛋白质的大部分功能也都不同。蛋白质图书馆和代谢图书馆一样，社区高度多样化。出于同样的原因，这种多样性使得庞大的基因型网络与探索蛋白质图书馆的过程相适应，蛋白分子在保存原有意义的同时，拥有进化成为功能不同的新蛋白质的巨大潜力。

代谢图书馆和蛋白质图书馆中充斥着基因型网络，这些网络由含义相同的文本构成，每个文本都被放置在高维空间的超几何体上，两个图书馆里的多样性社区数量也都多得难以想象。它们彼此间有诸多相似之处，但都与人类图书馆大相径庭。不过这也没什么好奇怪的，因为远在人类出现之前，它们就已经存在了。



确切一点说，代谢图书馆和蛋白质图书馆的出现至少比人类早了30亿年。从那时候起，蛋白质就从RNA手里接管了大部分生命的工作。这样做绝对有着充分的理由，因为蛋白质的构件要多得多，RNA只有4种核苷酸，相比之下蛋白质则有20个不同的氨基酸，大自然可以用蛋白质书写更多不同的文本。要写一条10个字母长的链，用4个字母的字母表大约可以写出100万条，而用20个字母的字母表则可以写出超过10万亿条，后者是前者的1　000万倍。相比RNA，蛋白质文本的数量要多得多，而且文本越长，两者的差异越明显。更多的文本意味着更多的构型，参与更多的反应催化，执行更多的功能和完成更多的任务。

但RNA的出现确实先于蛋白质，就凭这一点，RNA就足以在生物进化的万神殿里享有一席之地。如果没有历史上的第一个自我复制分子以及它的进化，也就没有今天的我们。不理解这种进化的过程，我们的工作也就不能算是完整的。

幸好，RNA和蛋白质之间有许多相似之处，这有助于我们理解RNA的进化。我们可以把RNA文本组织成一座超立方体图书馆，虽然不如蛋白质图书馆大，但依然规模惊人。在图书馆中，相似的文本离得近，相异的文本离得远。这座图书馆也属于高纬建筑，这意味着其中的社区比三维空间里的大得多，即一个文本附近有许多其他文本。由于RNA长链分子和蛋白质一样高度灵活，所以许多RNA文本的意义也会借助构型语言来表达。和蛋白质一样，RNA链也会在空间中弯曲扭转，精心折叠。

不幸的是，RNA与蛋白质的相似之处仅到此为止。RNA分子似乎不愿意轻易显山露水，对它的构型研究一直不顺利。科学家们至今只确认了数百个RNA构型，而我们已经知道了上千种蛋白质的构型和功能，相比之下，已知的RNA构型数量简直微不足道。我们在蛋白质中取得的成果，即大量比较自然界中的蛋白分子并绘出图书馆模型，暂时还不可能在RNA分子上重复。

尽管如此，得益于奥地利科学家彼得·舒斯特（Peter　Schuster）和他的同事们的工作，RNA图书馆的建设并不是毫无进展。舒斯特是欧洲计算生物学的鼻祖之一，他从20世纪70年代起就在维也纳大学担任教职，现在是该大学的退休教授。我与舒斯特的第一次会面似乎印证了许多欧洲人对奥地利人的固有印象。舒斯特个性活泼，块头很大，幽默感奇特，哪怕出现在奥匈帝国晚期的传统维也纳咖啡馆也毫不突兀，因为

那样的地方往往聚集了大量受过良好教育的饱学之士，从心理分析到量子理论，他们对每件事都能滔滔不绝。舒斯特正是那种传统的科学家，一个纯粹的知识分子，对各种话题都能信手拈来。他不摆架子，说话风趣，就算谈话内容再严肃也会找机会岔开话题幽默一把。有句耳熟能详的谚语阐释了奥地利人对生活本身和生活中诸多挑战的看法："情况不容乐观，但是绝望为时尚早。"舒斯特本人就是这句谚语的写照。

舒斯特看起来大大咧咧，但思维开阔、头脑清晰。他是提出RNA世界起源理论的先驱之一，他的研究小组开发的计算机程序成功预测了RNA一个重要的分子特性，即它具有二级结构。

当RNA单链进行折叠时，首先出现的正是它的二级结构。随着RNA链扭转、弯曲、缠卷，部分核苷酸互相配对，在分子内形成局部的双螺旋区段，像极了著名的DNA双螺旋。众多双螺旋与它们之间的单链区域，就构成了RNA的二级结构，而所有这些结构都是由一条RNA长链通过自身折叠形成的。RNA的二级结构就像蛋白质的α-螺旋和β-折叠，二级螺旋结构通过自组织形式，向更高级的立体结构自发发展。

舒斯特能通过核苷酸序列计算出RNA的二级结构，不仅如此，他们小组的计算机程序运行得飞快，几秒内就能预测出几百种分子构型。而对于更加复杂的RNA三维结构，我们至今都无法做到快速运算。有了如此高效的程序，我们就可以开始绘制RNA图书馆的地图了。即使我们离完全理解RNA的折叠方式和功能还有很长一段路要走，二级结构本身已经足够关键了：如果一个RNA分子的核苷酸顺序发生变异，破坏了它的二级结构，这个分子就不能进行正确的三维折叠。二级结构对RNA分子的意义至关重要，正如花束不能跳过一朵花凭空出现，RNA分子的三级结构也无法逾越二级结构存在。这个理由充分论证了研究二级结构的必要性。

舒斯特身边的研究人员在RNA图书馆中发现了许多可以解读的分子含义，令人眼花缭乱，而所有分子含义都可以通过RNA构型进行表达。举例来说，对于一个含有100个核苷酸的RNA，它的不同构型的可能数量已经达到了10^{23}种。许多天然RNA分子包含的核苷酸更多，更长的RNA链则相应拥有更多的可能构型。

另外，构型相同的文本在RNA图书馆中的组织方式和蛋白质图书馆非常相似，即文本连成网络，延伸至图书馆深处，你可以通过微小的改变积累，逐步修改，直到彻底改变一个文本的内容，同时保持文本编码的RNA功能不变。像蛋白质图书馆一样，RNA图书馆中不同的社区更像

中世纪的村庄那样多姿多彩，而不像城里的社区那样大同小异。每个社区里的构型都不尽相同，任意两个不同的社区之间的交集也廖廖无几。这一切都意味着RNA进化所遵循的规则与蛋白质相同，而最新的实验表明事实的确如此。

2000年，麻省理工大学的埃里克·舒尔特斯（Erik Schultes）和戴维·巴特尔（David　Bartel）完成了一个巧妙的实验，在RNA图书馆中开辟了一条新的道路。实验开始于两个较短的RNA文本，每个文本含有的核苷酸不超过100个。两个文本的许多核苷酸都不同，在图书馆中的位置相去甚远。不过它们并不是随随便便的RNA分子，两者都是核糖酶（ribozymes），是由RNA而不是蛋白质组成的生物催化剂。两个核糖酶扭曲形成不同的三维构型，负责催化不同的反应。第一个分子可以将一条RNA链分解成两条，第二个正好相反，主要通过原子键融合两条RNA链的末端，把两条链联结起来。我们姑且把这两种酶分别称作“切割酶”和“连接酶”。

如果你手头有一个切割酶分子，需要在图书馆某处找到一个连接酶，会是轻而易举还是难于登天？或者相反，通过手头的连接酶寻找切割酶呢？换句话说，以任意一个分子为起点，你能否通过对图书馆进行探索找到另一个想要的分子，就像分子进化能够实现的那样？如果你对基因型网络一无所知，就会认为这不可能，因为这两个分子天差地别。即便可能，大概也非常困难，因为一招不慎则满盘皆输，任何一点失误导致的缺陷分子在进化中几乎都意味着死亡。

舒尔特斯和巴特尔没有被困难吓倒。他们模仿自然选择的过程，从其中一个分子出发，一步一步改变核苷酸的序列，同时要求每一步都保留原本的分子功能，尝试向另一个分子靠近。他们利用自己掌握的化学知识预测穿过图书馆的可行路径，合成所有候选的变异RNA，测试新链能否像它的前辈一样催化相同的反应。如果不能，他们就尝试另一条不同的路径。

读到这里，你可能不会再对他们的发现感到惊讶了。从连接酶出发，他们一点一点地朝着切割酶靠拢，在这个过程中他们一共修改了40个核苷酸，同时保证新获得的RNA分子具有融合两条RNA链的能力；从切割酶出发，他们也以同样的方式，向着连接酶一点一点地修改核苷酸，在修改了40个核苷酸后，仍然保持新获得的RNA分子具有裂解RNA的能力。在两种分子相互转变的中间点附近，奇妙的事情发生了：只要对不到3个核苷酸分子进行修改，核酸酶的功能就会发生颠覆性的改

变。连接酶变成了切割酶，而切割酶则变成了连接酶。

这个实验像许多优秀的实验一样，传达了不止一条重要信息。第一条，具有切割酶和连接酶含义的文本不止一个，而是有很多。第二条，这些分子在RNA图书馆中相互连接，即便必须以保留分子功能为前提，你也可以循着这张网络找到功能相同的新分子。正是基因型网络使得这一切成为可能。第三条，循着网络中的某条道路前进，你总会在某个文本所在的社区中找到你想要的新性状。

在真实的进化历程中，探索RNA图书馆需要大量分子充当读者，而上述实验中只有一名孤单的读者。此外，科学家还要凭借生化知识指导这个读者，以防止它乱走，根据实验设计，它一直在沿着基因型网络前进。这使我心头一直萦绕着一个疑问：大自然中的RNA在进化上其实是随机而盲目的，对于现实中的进化而言，基因型网络是否对它也有助益？这个问题的答案直到10年之后，在我位于苏黎世的实验室完成一项与进化有关的实验后，才渐渐浮出水面。

大多数人认为进化非常缓慢，所需的时间与我们每个人的正常寿命根本不在一个数量级上。就人类的进化而言，这的确是事实，1 000年的时间才相当于50代人，但许多其他生物的世代时间就要短得多，比如大肠杆菌，它每20分钟就能繁衍出下一代。繁殖50代大肠杆菌甚至用不了一天。一个RNA分子几秒钟就能完成自我复制，RNA使用的分子复制体系与复制DNA的那一套相仿。不消一天，你就能得到上千代RNA。

有了快速复制的生物体和分子，实验室就可以开展一个雄心勃勃的实验：重演进化。类似的模拟进化实验能够让科学家形象地看到进化如何在生物传宗接代的过程中，逐渐改变整个生物种群。由于RNA分子对早期生命至关重要，它们在这种实验中就显得特别有吸引力。RNA分子兼具自我复制和变异的遗传特性，本身的性状又能够作为自然选择的作用对象，集各种进化的要素于一身。

2008年秋，我为开展实验室里的进化实验面试了一些新人，希望能解决萦绕我心头多年的疑问。其中一个年轻的美国科学家非常出挑，不仅因为他的证书，也因为面试时他穿着徒步鞋。虽说大多数学院的科学家都崇尚休闲打扮，对其他职业严格的着装要求往往嗤之以鼻，但徒步鞋还是有点不同寻常。起码，这显示出了他的健康自信。

这个年轻人叫埃里克·海登（Eric Hayden），是一名化学家，刚从俄勒冈州立大学取得博士学位，他在RNA酶方面进行的工作非常杰出。

虽然之前对进化生物学所知甚少，但他对进化表现出了深深的好奇。他面容坦率，一个微笑就照亮了房间，顿时让人心生好感。我和他简短地谈了谈，随后就让他和组里的其他研究员随便聊聊，看看他在我们的圈子里是否自在。他肯定觉得就像在家里一样舒坦：一个小时后回到我办公室时，他已经脱下靴子只剩袜子了，他解释说靴子太热了。

埃里克得到了这份工作，而我从未后悔这个决定。他非常了解RNA，做实验谨慎仔细，非常讨人喜欢。我觉得能和他一起工作是我的荣幸。

埃里克在我的研究小组里负责研究一种核糖酶，这是一种RNA酶，用以帮助某些细菌表达基因。这种酶可以根据特定的核酸序列辨识相应的RNA链，然后对目标RNA进行切割，随后将自身连接到其中一个片段上。（许多生物细胞内都有能识别并裂解特定DNA和RNA的分子，它们的目的多种多样，例如破坏入侵病毒的外来DNA，或者把较短的DNA片段连接成相对较长的DNA。）关于这种酶，我只想知道基因型网络能否改变这种酶，使它能够识别新的RNA分子。

为了找出答案，埃里克把这种酶复制了10亿多份——所有复制所得的酶正好相当于一汤匙，然后用分子复制体系对这10亿个酶进行扩增。生物本身的复制体系并不完美，偶尔会出现复制错误，因此获得的少量核糖酶会出现变异。随后，埃里克运用某种化学方法，设法让变异体中那些依旧能够识别原目标RNA的核酸酶继续复制。这种手段巧妙地模拟了自然选择的核心要求：只有保留原先功能的变异分子才能在生物的代际之间传递。

埃里克的实验循环了许多次。第一次循环前，所有的分子都一模一样，就像10亿个读者在图书馆里弯腰读同一卷书。第一代分子诞生后，许多分子已经变异了，其中只有一部分能活下来。幸存的分子在衍生出第二代分子时进一步发生变异，循环往复。仅仅过了10代，子代与初始RNA的平均核苷酸差异就已经达到了5个，有些甚至相差多达10个。这10亿个读者已经在图书馆里一哄而散了。

只需要上述简单的观察，我们就可以得出埃里克实验的第一个结论：经过数代复制，这群RNA分子已经不是完全相同的了，它们与初始RNA分子之间多少存在数个核苷酸的差异。尽管基因型已经改变，选择机制却使得所有分子保留了相同的功能特征。由于不同的RNA分子分布于图书馆的不同位置，埃里克的实验已经表明，基因型网络在RNA图书馆中同样存在。

实验的第二个结论涉及两类分子。第一类分子是我们在上面所说的“混合”分子，而第二类分子则是初始RNA分子，其中每个分子都一模一样。埃里克让两类分子切割一种新的RNA。他用一个硫原子取代了原先的RNA链上的一个磷原子，如此一来，酶的工作艰难了许多。埃里克用复制-选择的方式分别处理两组分子，让它们各自经历模拟的进化过程。但他现在只选择能够切割新RNA链的分子，接下来他想知道，四散在基因型网络上的第一类分子和集中在基因型网络一处的第二类分子，到底哪类分子上手新任务的速度更快。

如果基因型网络有助于分子进化，那么第一类分子应该做得更好，因为第一类分子能在RNA图书馆里探索更多社区。这正是埃里克的发现。他在多样的混合分子里发现了一个RNA酶，这个核酸酶执行新任务的效率比单一的第二类分子中的核酸酶高8倍。

埃里克的实验还有一个额外的惊喜。我们是在测定那个效率奇高的新分子的序列时意外发现的。

很多研究人员已经研究过我们实验开始时用的RNA酶。这个分子很小，只有大约200个核苷酸。我们知道相关核苷酸序列、折叠方式、相关功能以及作用原理。大部分你可能想了解的内容我们都知道。在酶分子的数轮复制中，我们都精确控制了分子进化的环境，甚至精确到每个分子的浓度。既然对全局有了近乎周详的掌控，你可能会觉得我们应该能预测分子会做出怎样的改变来适应新任务。就像如果你彻底了解一架机器，了解机器上的每一个齿轮、每一个螺钉、每一根杠杆、每一根弹簧以及它们的协作方式，你肯定也知道改进这架机器的最佳办法。

但是我们不知道。我们完全无法预测大自然究竟会如何改进实验中的酶分子。直到今天，我们仍然没有完全理解为什么这样改进的效果最好。

实验室里模拟进化的实验常常出现意外。不论我们对一个分子研究得多透彻，不论实验多简单，不论控制得多精确，自然总是出人意料。哪怕是最简单的酶，也比大多数人类制造的机器要复杂难懂得多。

虽然我们在预测最佳结论上毫无建树，但也没有空手而归，我们现在已经知道基因型网络可以加速生物种群进化的速度。这个结论正中要害，虽然我们无法预测某个个体的新性状，但这并不妨碍我们在种群层面上对于进化的研究。

科学祛魅自然，确定自然法则，剥夺人对世界的惊奇和敬畏之心，

这让很多伪科学人士深感困扰。用诗人约翰·济慈（John Keats）的话来说，科学家是群扫兴的人，“使天使折翼”（clip an Angel's wings）、“拆开彩虹”（unweave a rainbow）。达尔文理论之所以不被接受，这种情感当然也是原因之一，不过上述实验表明，我们依然可以想到两全其美的办法。科学能够解释进化的普遍原则，但是不能预测单个进化。理解进化的能力丝毫不会影响进化的魔力。这本身就是我们对自然保持惊奇和敬畏的理由。



ARRIVAL OF THE FITTEST

05 命令与操控

Solving Evolution's Greatest Puzzle

无论多复杂的生物，它的形态和功能都受到调节因子的控制。调节因子占据着某个基因相邻的一小段 DNA，一旦它们遇上特定的 DNA 序列，就会与之结合。调节因子与相应的 DNA 需要在形态上互补。有些基因表达能被调节因子抑制，有些基因则需要它们激活。调节因子指导着所有生物的发育。调节因子之间相互调控，形成了复杂的网络。



一直以来，牛奶都被当作一个正面的意象。麦克白夫人的丈夫在犯下弑君罪行前，因为心中尚存“纯良人性的乳香”而犹豫不决，《出埃及记》的第3章向希伯来人承诺了“遍地流淌着牛奶和蜂蜜的理想地”，直到今天，我们还用“如母乳般纯良”来形容无害的事物。然而，对于这个世界上超过半数的人来说，一杯牛奶并不意味着健康的生活，而是如假包换的毒药。牛奶对他们来说意味着胀气、放屁和腹泻，这是因为他们体内缺乏消化牛奶中特有的乳糖的消化酶。没有这种酶，人体就无法降解乳糖，而人体内的微生物很乐意清理原封未动的乳糖，正是它们的代谢产物引发了人体的不适反应。

乳糖不耐的人曾经也能够消化母亲的乳汁。在幼年时，他们体内的乳糖酶基因是激活的。用专业的术语来说，这些基因是表达的（expressed），基因表达的意思是：编码乳糖酶的DNA指令被转录为RNA，RNA继而被翻译为相应的蛋白质，也就是酶。乳糖不耐就是源于成年人体内的乳糖酶基因被永久关闭，不再表达。这种可以激活或者关闭的基因，我们称之为“可诱导基因”（regulated genes）。

对于多数人来说，成年后体内的乳糖基因关闭才是常态。如果你有幸能够耐受乳糖，那么说明你在乳糖酶基因控制区存在一个突变，这个紧邻酶基因的突变使得你的乳糖酶基因在成年后仍然可以持续表达。由于这种耐受乳糖的突变最早广泛流行于从事畜牧业的人口中，所以没准你的某个遥远的祖先就是一位爱喝牛奶的奶农，悠闲地生活在东非或者斯堪的纳维亚。乳糖耐受能力如野火燎原，只是眨眼之间——人类发展畜牧业的历史只有大约区区8 000年，在某些人口中的突变率就从0%蹿到了90%。乳糖耐受是近世代自然选择在人类基因组中留下的最深刻的烙印之一。

说来可能没人相信，但是乳糖诱导的消化不良与自然进化有着密切的关联。两者的联系在于调节，类似于乳糖酶基因开关的分子调节。除了引起肠胃不适，基因调节还与数不清的生物形态有关，如水母波动起伏的“伞”，鲨鱼犹如水雷般致命的身形，玫瑰窈窕纤细的茎秆，红杉树巨大粗壮的树干，毒蛇吓人的条形躯干，野兔疾步如飞的四肢，还有鸟儿用以翱翔的双翅。从细胞中第一个平衡细胞的生长开始，基因调节就从细胞依旧利用RNA作为基因组的远古时代出现了。30亿年之后，地球上每一种生物形体的发育和塑造中都有基因调节的参与。如果我们不能理解新的基因调节如何出现，也就无法完全理解新性状的进化如何完

成。

虽然无论生物多复杂，它们的形态和功能都受到基因调节的控制，但是对其的研究最容易在简单的单细胞生物中开展，比如细菌。两名法国遗传学家，弗朗索瓦·雅各布（Francois Jacob）和雅克·莫诺（Jacques Monod）正是借此获得了诺贝尔奖。他们的工作开始于20世纪50年代，当时DNA双螺旋模型刚面世不久。他们主要阐明了原始的细菌，比如大肠杆菌，如何通过调节基因的表达使自身获得代谢乳糖的能力。

基因表达始于一种复制分子，我们在第4章埃里克·海登的实验中简单介绍过类似的连接酶。这种复制分子是一种聚合酶，能够催化合成聚合物，也就是由许多更小的单位构成的长链分子。在基因转录为RNA的过程中，这些基本的单位分子就是4种不同的核苷酸。当RNA聚合酶要转录一个基因时，它会首先黏附到目标基因的DNA上，并沿着DNA序列一个碱基一个碱基地滑动，同时合成一条RNA链，它的碱基对序列和目标基因的完全对应。细菌正是通过这种方式合成了它们的乳糖酶变体，我们称之为β-半乳糖甘酶（beta-galactosidase，这个酶的名字非常拗口，所以通常会缩写为betagal）。它会把乳糖分解为结构相对简单的葡萄糖和半乳糖，而后其他酶再利用这两种糖摄取所需的能量和碳原子。

细菌以一种称为转录调节因子（transcriptional regulator）的分子控制β-半乳糖甘酶基因的转录。通常情况下，调节因子的作用只有一个：它占据着某个基因相邻的一小段DNA。细胞内的液态环境中漂浮游荡着无数这样的调控分子，一旦它们遇上特定的DNA序列，即一段DNA“指令”，就会与之结合。不同的调节因子识别的序列也不同，β-半乳糖甘酶的调节因子识别的序列为G-A-A-T-T-G-T-G-A-G-C。

与酶一样，让这种识别成为可能的同样是依靠蛋白质分子的空间折叠。调节因子与相应的DNA需要在形态上互补，就像能够互相拼接的乐高积木。这个比喻很形象，但是并不太贴切，因为性状并不是互补的关键。确切地说，相互靠近的两个分子必须发生相应的形变，否则就无法发生互补。另外，乐高积木只有10多种不同的形状，而分子的形状则丰富得多，蛋白质有数万种不同的结构，而DNA的结构数量则更在这个之上，几乎和人类语言中所有的词汇数量相当。

除此之外，与乐高积木不同，许多分子的形状改变是自发的，不仅像酶一样发生在平时的分子震动中，同时也发生在分子间相互结合的时候。这种形变就像你用正确的钥匙开锁：只有在正确的钥匙插入的时候锁芯才会转动，门才会打开，只不过在分子中，是热能而不是钥匙在转

动“锁芯”。

像乐高积木一样的调节因子在以一种最简单的方式调整β-半乳糖甘酶的合成：它识别的关键词恰好在多聚酶转录的起始位置，调节因子与之结合抑制了多聚酶的作用（如图5-1的上半部分所示）。当环境中没有乳糖分子存在时，调节因子（R）与“关键词”结合并阻止多聚酶（P）读取基因序列，于是基因就处于关闭状态。

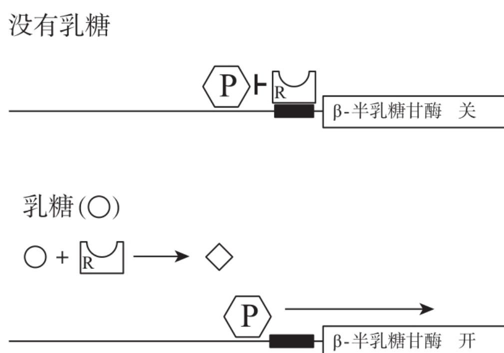


图5-1　基因调节

为了能够利用乳糖，当环境中出现乳糖时，细菌必须找到去除转录障碍物的方法。如果说调节因子不仅能够与特定序列的DNA结合，还能与其他分子结合，就像一块乐高积木可以与其他许多积木拼接，那么细菌利用乳糖的能力就更好理解了。所谓的其他分子，正是乳糖。当作为钥匙的乳糖插入到作为锁的调节因子里后，后者的分子形状就会发生改变（如图5-1中的菱形）。产生形变的调节因子从DNA上脱落，这时多聚酶就能自由地进行转录，一个接一个碱基将其转录为RNA，然后再经过细胞翻译，量产出β-半乳糖甘酶。总而言之，只要周围的环境中还有乳糖，β-半乳糖甘酶基因就可以被激活，继而被合成，否则由于转录受到阻碍，乳糖酶基因将再次关闭。

β-半乳糖甘酶是一种了不起的酶，但“天下没有免费的午餐”。一个表达β-半乳糖甘酶的细菌体内可不会只有几个β-半乳糖甘酶蛋白质，而往往会有3 000多个β-半乳糖甘酶分子，合成它们中的每一个都需要超过1　000个氨基酸分子，合成的原料和能量都需要细胞提供。按照常识，细胞应当对β-半乳糖甘酶的合成进行调节以避免浪费原料，不过我们不能仅凭直觉揣测大自然的想法，不然生物学家们可就都要失业了。由于同一时间内有细胞在同时合成数百万个不同的分子，所以持续合成β-半乳糖甘酶的消耗很可能微不足道。此外，让基因长期处于激活状态还有

一个好处，那就是当环境中突然出现乳糖时，细菌能够在代谢上抢得先机。

2005年，以色列魏茨曼研究所（Weizmann Institute）的埃雷兹·德克尔（Erez Dekel）和尤里·阿龙（Uri Alon）希望能够弄清表达β-半乳糖甘酶的确切消耗。他们欺骗细胞，让它们以为周围的环境中有乳糖，实际上却并没有。即便如此，细胞依旧激活了β-半乳糖甘酶基因，如果这种浪费足够显著，它将会在细胞的分裂速度中有所体现。事实上也的确如此，它们的分裂速度降低了数个百分点。打个比方，这就像资金周转不畅的开发商在房屋施工的时候，非要修一个他并不需要的游泳池，游泳池占用了他的资金和物料，最后只能牺牲室内的装潢。相比之下，另一个更优秀的建筑商会尽快完工，卖掉房子之后再建新的房子，而此时上面说到的那个开发商还在为游泳池里铺什么样的瓷砖而头疼。

仅仅几个百分点的工程拖欠似乎算不上什么大事，对于大肠杆菌20分钟左右产生一代的分裂速度而言，一分钟的差距好像不足为奇。但是这一分钟的延迟从长远来看却是致命的。如果一个菌群中有50%的细菌存在这一分钟的缺陷，80天之后，存在缺陷的细菌数量将不足1%，而300天之后，这个比例会降到百万分之一以下。它们很快就会不可避免地被繁殖相对较快的同类排斥殆尽。自然选择向来雷厉风行，不讲人情。

如果调节能够避免不必要的浪费，那么它应当无处不在。事实上也的确如此。想象一下，一个包含数百种生化反应的代谢，如同数百条互相连通的管道，而乳糖代谢只是其中之一。营养物质流入管道，而流出的则是生物质。每一条管道都有一个专属的水泵，作为水泵的酶分子会推动原料分子通过管道，细胞能够根据自己的需要调节每一个水泵的工作。如果细菌在土壤里发现了新的食物，比如一个掉落的苹果或一具腐烂的尸体，它们就会打开对应水管里的水泵。一旦营养物质消耗殆尽，水泵就会被关闭。此外，如果环境中某些营养物质的供应增加或减少，细菌还能够将水泵的速度调节到恰当的大小。

β-半乳糖甘酶的基因表达能够被调节因子抑制，而其他基因的调节方式则正好相反：这些基因平时也处于关闭状态，只有在需要的时候才会被激活，即它们的调节因子帮助基因在需要的时候进行转录，而在不需要的时候抑制多聚酶结合。虽然转录水平的调节是所有调节中最重要的，但它并不是唯一的调节方式。细胞还能够调节RNA翻译成蛋白质的速度、蛋白质的活性、蛋白质的寿命等。调节方式的多样性大概最能够

用来说明调节本身的重要性：生命会在10多种不同的水平上进行调节。



我们在这里想象一家高档餐厅的厨房，厨房的食品储藏室里整齐地摆放着各种蔬菜、肉类、鱼、食用油、香料以及调味料。这些食材足以用来烹制任何你想吃的菜品，从家常小炒到异国风味，每一种都色香味俱全，厨师长还要求厨房能够24小时进行供应。调节因子在细胞内扮演的角色之一就像这个厨房里的抠门经理，勒令厨师只能按需取用食材，不忍心把钱浪费在哪怕一个额外的马铃薯上。

不过调节因子扮演的角色不只是经理，还同时兼任厨师长，手握决定每天食谱的大权，指导其他人哪里应该加入一量杯的豆子，哪里应该添加两量杯的鸡汤，哪里又需要少许盐，然后用350摄氏度的烤箱烤制30分钟。这里所说的食谱，正是基因组中每一个基因精密的表达方式。基因的正确表达可以使细胞内每一种蛋白质的数量都保持在恰当的水平。

考虑到生命的复杂性，把一条蓝鲸的基因表达和一个蛋奶酥的配方相提并论，多少显得有失公允。任何一种细胞内的蛋白质成分都远比最精致的料理复杂，数千种蛋白质分子的数量和合成时机在细胞内受到精确调控，哪怕技艺最精湛的五星厨师都对这种火候的控制望尘莫及。不仅如此，进化还在孜孜不倦地研究着新的“菜色”，细胞、组织、器官乃至整体的新性状，都是不断变化的、庞杂的调控系统的产物。

生物调控是发育生物学研究的议题，发育生物学是生物学中研究一个细胞如何发育为一个生物整体的分支学科。发育的过程十分神奇。发育生物学试图解释生物体内的细胞为何不仅仅是一坨松散无形的囊泡，而是能在动物体内发育出如心、肝、肺、脑等器官，在植物体内发育出根、茎、叶、花等构成。

每种器官都有高度精专的分工，并含有许多特异的细胞种类。以你的心脏为例，其中的细胞就包括泵血的心肌细胞、将心肌细胞联络在一起的结缔组织细胞，以及像龙舟上给桨手们提供挥桨节奏的鼓手那样，通过电信号控制心脏搏动节律的起搏细胞，这些都是心脏特有的细胞类型。那么这些特异的细胞是如何从同一个受精卵分化而来，又如何在恰当的时间和位置发生分化的呢？一个细胞要如何知道自己应当分化成起搏细胞，而不是一个神经元或者干细胞呢？

答案就是调控，调控指导着所有生物的发育。多细胞生物体内的细胞通过合成特异的蛋白质完成相应的分化。我们体内的每个细胞都包含有人类全部的基因，细胞的区别源于它们选择性表达的基因。肌肉细胞能够表达马达蛋白，这种小小的分子机器是肌肉细胞能够收缩的关键，所以几乎所有种类的肌肉细胞都表达这种蛋白质。人类的眼睛内有一种透明蛋白，能够透光并将光聚焦到感光的视网膜上。软骨细胞能够表达胶原蛋白和弹性蛋白，作为缓冲物以防止关节骨之间过度的摩擦和损耗。

已分化细胞和特异蛋白之间的关联并不简单，虽然不同的分化细胞的确各自表达着独特的蛋白质，但蛋白质并不能代表细胞的种类。实际上，任何蛋白质都会在多种细胞中表达。眼睛内的玻璃体，即位于角膜和视网膜之间的透明胶体，其中的胶原蛋白与软骨细胞合成的无异；肱二头肌肌肉细胞与心肌细胞合成的马达蛋白同样别无二致，类似的例子不胜枚举。决定一个细胞“身份”的不是某一种独特的分子，而是分子指纹（inolecular fingerprint），即一个细胞内所含有的数百种蛋白质的组合方式。所以新的细胞种类就意味着新的分子指纹，也就是调控下的基因表达的新形式。

对细胞分化起关键作用的基因往往在许多不同类型的细胞中都能表达，所以对这些基因的调控往往需要多个开关。在图5-2中这些开关以小的矩形方框表示，每一个矩形的方框都代表一段不同的关键词，每个关键词都有与之结合的调节分子（图中的其他形状）。典型的例子有编码晶状体蛋白的基因，这是眼内晶状体分子指纹中的成员蛋白之一，正是依靠它我们的眼睛才能进行聚焦（我们会在第6章对这个基因展开更多讨论）。

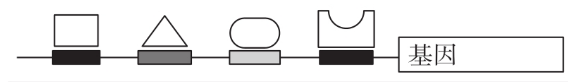


图5-2　一个基因与多个调控分子的对应关系

晶状体蛋白至少有5个调节分子，Pax6就是其中之一，它通过结合在基因附近的区域决定基因表达与否。有的调节因子与DNA结合紧密，所以能够强烈影响基因的表达；而有的则结合疏松，对转录的影响也就相对较弱。调节分子对多聚酶转录基因的干预，就如同内阁议员们向国王进谏施压。有的调节分子倾向于抑制基因表达，有的则倾向于激活；有的对基因表达影响重大，有的则无足轻重，所有调节因子的效应总和

决定了基因表达与否。

那么是什么在调控调节因子？很简单：其他调节因子。图5-2中调控基因的所有调节因子本质上都是蛋白质，和其他所有蛋白质一样，它们都由各自的基因编码，而基因则受到调节因子的调控。调节晶体蛋白表达的Pax6不仅在晶状体内，也同样在角膜、胰脏以及发育中的神经系统内表达，它的表达受到多种调节因子的共同调控。那么如何调控这些调节因子呢？当然是再依靠其他调节因子。那么调节因子的调节因子呢？当然是依靠新的调节因子。所有这些调节因子形成了一条花环链，调节分子之间的级联关系如图5-3所示。

一眼看去，图5-3中的级联关系已经足够繁杂了，不过现实中的基因调节远比这复杂得多：调节因子之间的相互调控不仅是线性的，甚至可以是环形的。图5-4示意了5种调节因子基因之间形成的环状调控通路，5种基因依次用小方块和方块里的A到E表示。出于简便考虑，图里没有画出调节因子在DNA上的识别位置，只标出了调节因子之间相互调控的关系，黑色的箭头意味着调节因子能够激活目标基因，而灰色的直线则表示调控因子会抑制目标基因。简而言之，这些基因之间能够相互促进或者抑制。图中的虚线则表示情况更加复杂的关系：每种调节因子还掌控着其他数百个基因的命门，它们都不在这个环里。

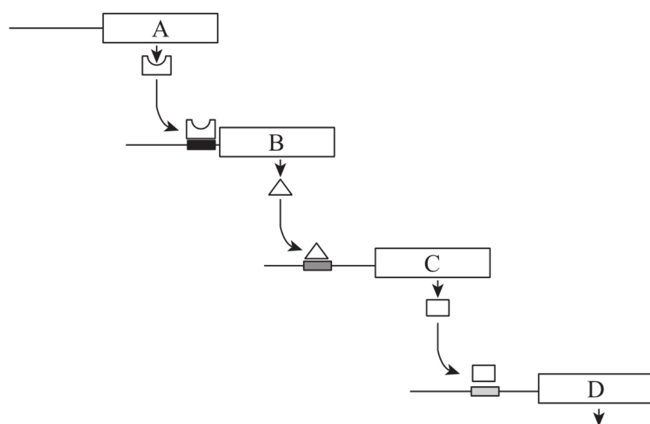


图5-3　调节因子级联图

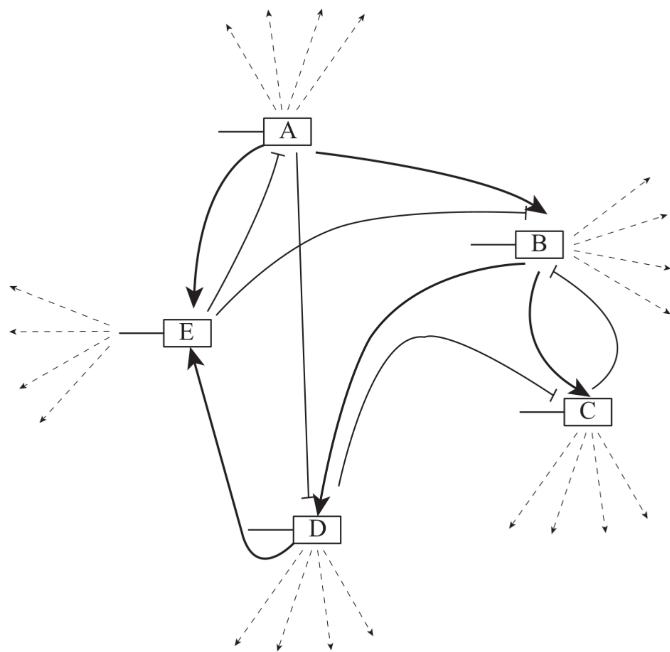


图5-4　环形调控回路

Pax6基因正是这种环形回路里的一员，它的变异会引起严重的后果，也反映出调节环路的威力：人类具有Pax6基因缺陷会导致先天性虹膜缺失，同时伴有晶状体浑浊以及视网膜退化，从而致盲。作用与Pax6类似的基因在许多动物身上都扮演着相同的角色，包括老鼠、鱼类和果蝇，虽然它们的眼睛与人类相比在结构上有着天差地别。果蝇体内的“Pax6”被称为eyeless，顾名思义，没有eyeless基因的果蝇无法发育出眼睛。不过当果蝇体内的eyeless过剩时，结果更惊人。生物学家在果蝇胚胎内原本不会表达eyeless的部位激活了这个基因的表达，结果是果蝇在触角、腿部甚至翅膀上都发育出了眼睛。

图5-4看起来有点像工程师画的布线图。考虑到基因的不寻常性，即便基因之间并没有线条把它们连接起来，这样的比喻也还算合理。类似的布线图是一种记录环路基因型的简洁手段，其中包含编码调节因子的DNA以及调节因子识别与结合的关键词信息。只需要简单一瞥，你就能知道图5-4中的基因A能够激活基因B和C，而D能够抑制C等。在活细胞内，基因之间的相互激活和抑制构成了一部交响乐，每一个基因都相当于一种乐器，它们跟随着相互之间的旋律与节奏演奏，直到整个环路达到平衡——就像复调闭和弦，环路中所有基因的表达都不再变化。

在这个平衡点上，环路内有的基因被关闭，而有的则被激活。举个

假想的例子，图5-4中的A和C基因可能在平衡后处于激活状态，而B、D和E则会被关闭。所有基因的开闭状态（例如，“开”“关”“开”“关”“关”）被称为“基因表达谱”（gene expression pattern），由于环路里的基因调控着许多其他基因的表达，所以基因表达谱除了是环路本身的表现型，同时也决定了细胞的分子指纹。基因表达谱是又一种无法被直接感知，只能通过精密设备进行测量的指标。但它又与最明显的表现型有关，即生物躯体的形态。于是，想要新的生物形态首先要有新的基因表达谱。

基因调节环路塑造了千奇百怪的生物形态，从聚集在腐烂水果上的果蝇到遍地丛生的拟南芥，再到斑马鱼——一种身长不到10厘米、全身布满条纹的淡水鱼。这几种生物都非常不起眼，但是有两个特征使它们成了研究发育的理想实验对象：它们身形娇小且繁殖迅速，能够让我们在短时间内研究大量的个体样本。

我们从它们身上学到的一点是，调节环路对身体形态的调节速度非常快，令人难以置信。黑腹果蝇（drosophila melanogaster）的幼虫在果蝇产卵后的15小时之内就会孵化，紧接着在7天之内就会完成化蛹和变态，发育为成年果蝇。一个生命从无到有只需要15个小时，还可以独立捕食、爬行、在世间游荡，这也就难怪成千上万的科学家殚精竭虑想要弄清楚果蝇的基因到底是如何工作的。

果蝇的身体主要有三个组成部分，这三部分又可以再细分为14段体节：头部为单独的一部分，胸部占三段体节，腹部占11段，每一段体节各司其职，负责爬行或是生殖的功能。对于大多数人而言，果蝇既不漂亮也不高贵：这种卑微的虫子根本无法与翅膀华丽的鸟儿，抑或与雄伟的巨型红衫相提并论。不过，14段体节和它们的作用对于果蝇来说，犹如哥特式教堂的飞扶壁和帕提侬神庙的多立克柱式，科学家研究体节得到的启迪堪比在生命科学任何其他领域的所得。直到今天仍然有许多人，从学习生物学的高中生到诺贝尔奖得主，都还在研究果蝇的体节。体节是研究基因调控的理想素材，体节研究中得到的许多原理在其他动物体内同样适用。

果蝇在产卵前会在卵内植入一些短小的遗传物质片段作为化学信号，帮助幼虫发育。这之中就包括一种名为bicoid的基因的RNA转录产物，利用它，果蝇卵能够翻译出一种bicoid蛋白。（没错，研究果蝇的生物学家在取名的时候通常没那么讲究。）bicoid蛋白集中在受精卵的前部和将来会发育为果蝇头部的位置。bicoid蛋白就像一滴坠入水里还来不及扩散的糖浆，它在果蝇受精卵内的浓度在离开前端之后呈现迅速

衰减态势。

除了bicoid之外，果蝇妈妈还在受精卵的前端留下了几种其他基因的RNA转录产物，同样的道理，受精卵的后端也有独特的化学信号，在离开后端之后它们的浓度也迅速衰减。果蝇妈妈的工作完成之后，胚胎的每一个部分都会拥有自己独特的调节因子组合，如同条形码一样，独一无二。

当一个精子与卵子相遇，受精卵便会形成并开始分裂。胚胎发育中会根据母亲留下的RNA分子决定合成蛋白的种类。在每一个胚胎子细胞内，蛋白质的合成数量都由母亲预留的RNA决定。而这些胚胎中早期合成的蛋白质正是决定其他基因开闭的调节因子，某个基因的开或关取决于对应调节因子的数量。举个例子，如果一个基因的激活因子只在受精卵的前端十分丰富，如bicoid，那么这个基因只会在受精卵前部被激活表达。

在早期调节因子调控的基因中有一些比较特殊，它们本身也是调节因子，用来激活别的基因，而进一步激活的基因中又有一些是编码调节因子的，以此类推。不仅如此，调节因子之间还会逐渐形成相互调控。最复杂的调控环路中涉及多达15种不同的基因，这个环路里的基因均执行着我上面所说的相互调节，结果是有的基因最终得以表达，而有的则没有，形成了自己独特的基因表达谱。

调控环路中有一种格外特殊的蛋白质，名叫齿状蛋白（engrailed）。经过与别的基因相互作用和影响，齿状蛋白在胚胎里呈现高度规律的间隔表达。果蝇胚胎中有7个区域表达齿状蛋白，另外7个则不表达，两种区域间隔分布，形成了果蝇体节最早的划分依据。接着，齿状蛋白以及别的调节因子继而控制其他基因表达，决定一段体节究竟是发育出腿，还是萌出翅膀，抑或是参与构成腹部。

上面的过程以及更多没有提到的变化都发生在数个小时之内。不过果蝇胚胎受到发育生物学家们的普遍青睐，不仅仅是因为它的发育速度：在果蝇体节完全形成之前，胚胎细胞之间还没有被彻底分隔。换句话说，分子在成长的胚胎里能够自由出入不同的细胞。而对于大多数其他物种而言，胚胎细胞在受精后几乎会立刻与别的细胞分隔，这让细胞间的交流通信变得异常困难。

当然这并不意味着根本不可能。男性的生殖器官阴茎和阴囊就是一个典型的例子。当男性胎儿发育到8周的时候，一小簇睾丸间质细胞

（leydigcells）就会在将来发育出性器官的位置附近释放雄性激素。雄性激素中包括睾酮等对塑造性器官至关重要的激素，激素会指导周遭的细胞向阴茎和阴囊分化，并在日后分化出精子细胞。雄性激素从睾丸间质细胞内被分泌出来后会进入细胞之间的空隙，雄性激素的分子结构能够让它随意穿过细胞膜，进入另一个细胞内。

在新细胞内，雄性激素的受体，即一种能够识别雄性激素分子形状的特殊蛋白质，早已等候多时。当两者相遇时，受体分子的形状就会发生改变，形成一把分子锁。蛋白质的形变让它能够识别DNA上的某个关键词，并激活相邻的基因。雄性激素受体能够激活许多不同的基因，其中就包括某些调节因子，它们在雄性器官内维持着数百个基因的激活状态，正是这些基因赋予了男性生殖器官中的细胞独特的分化性质。

从果蝇到人类，胚胎发育的每时每刻，所有组织内都在发生类似的信号交联，涉及的信号分子数以百计。正是在这种超乎常人想象的信号交流过程中，细胞得以确定自己的位置和命运，就像表达bicoid的细胞们“知道”自己位于胚胎的“头等舱”一样。基于同样的原理，细胞在信号指令的操控下分裂、移动、膨胀、收缩并变得扁平，最终完成细胞分化和生物塑形。不管何时，当细胞需要发生分化，生物形态需要进行重塑时，都逃不过细胞对信号分子表达的调整。

如果我们能够弄清从果蝇到人类胚胎发育的调节方式，我们就能预测器官、组织和细胞的形成，以及为何不同的生物在外形上千差万别。如此，真可谓大功一件。然而不幸的是，环路体系的表达谱着实庞杂，即便像图5-4中画出的经过简化的环路依旧十分复杂。如果说A能够激活B，而C却抑制B，B能够激活C，而D则抑制C，那么我们很难一眼看出各个基因最终的表达情况。实际情况是，许多现实中的环路含有的基因数量比图5-4要多得多，数十种调节因子像尼龙绳一样相互交织，繁复程度远远超过我们大脑的处理能力。不过也不是毫无办法，我们还有能够利用数学运算模拟环路内分子关系的计算机，与我们的碳基大脑不同，科学家可以依靠硅基大脑的算法，预测环路内所有基因最终的基因表达谱。

曾经有一名杰出的计算机科学家耗费毕生精力试图完成这项工作，他的名字叫约翰·瑞尼茨（John Reinitz）。20世纪90年代，当我还是耶鲁大学一名研究生的时候见过约翰。他比我年长几岁，大家都称他为怪胎，在那个抽烟并不光彩的年代，他时刻烟不离手。即便是在星期五便装日，他也衣冠楚楚，一丝不苟。他开着一辆古董级的大众甲壳虫，后

座上堆满了垃圾食品的包装盒。约翰叛逆、不拘小节，他敢于挑战主流的勇气对他的研究来说简直是无价之宝。

当时有许多科学家在研究果蝇的胚胎，而计算机对于他们的价值仅限于写论文。多数人研究果蝇主要是通过改变DNA与编码的基因，或者是在实验室里控制某种调节因子的表达，然后观察这些改变对于体节发育的影响。这些科学家的实验同样多产：其他的暂且不论，他们在果蝇基因组中找出了数千种与胚胎发育有关的关键基因。但是对于理解整个表达环路中的基因表达谱而言，单个基因在整体中显得微不足道，实验科学家们一次只能针对一个基因的研究手段注定收效甚微。虽然今天科学界已经普遍接受了这项技术，但是在20世纪90年代早期，约翰试图用计算机模拟果蝇的想法根本不被一众科学家看好，甚至遭到了少数人的无视及鄙夷。

约翰的想法有点像建造一台飞行模拟器，后者对于培训空军和商业飞行员来说不可或缺，它不仅可以模拟整套驾驶舱的操作机械，还能够模拟飞行中受到的气流干扰及仪表故障。与之类似，约翰收集了果蝇胚胎早期发育中的各种调节因子，以及它们相互之间调节关系的海量信息，将这些信息代入算法，并在计算机中模拟果蝇发育的过程。就像那些运行效果优良的飞行模拟器一样，约翰的果蝇模拟器也得以顺利运行——这可不是一件容易的事。果蝇模拟软件能够模拟果蝇胚胎的早期发育，而且运算速度惊人。它能够不断重复运算，直到保证没有任何算法遗漏。正如飞行模拟器能够模拟坠机，除了演算正常胚胎的发育，果蝇模拟器还能模拟基因表达调节异常的情况下，不同基因突变如何导致胚胎发育畸形。

我在这里写的这几行字几乎相当于约翰在过去几十年里花费的全部心血，他在建立果蝇模拟器的过程中受尽了同行的冷漠和蔑视。当我抬手准备拍死一只苍蝇的时候，脑海里经常会闪过他默默奋斗的身影。（然后世界上就少了一只苍蝇。）



除了脊梁骨和脊髓，包括鱼类、哺乳类、两栖类、爬行类和鸟类在内的6万多种脊椎动物在形态上可谓千姿百态。不过，由于所有的脊椎动物都可以追溯到生活在5亿多年前的同一个祖先，所以我们可以说现代脊椎动物的形态多样性都建立在类似的内部结构之上。比如偶鳍，通

常一对在躯干前方，一对在躯干后方，用以帮助鱼类在水中推动身体前进并控制前进的方向，陆生动物用于爬行奔跳的四肢正是起源于偶鳍。而某些陆生动物的前肢后来又进化为鸟类的翅膀，比如恐龙。

四肢是陆生脊椎动物进化的关键所在。无论前后肢，都是由结构类似的三个部分构成的，上臂和大腿，前臂和小腿，以及手和脚。人类的手臂与腿部的主要骨骼与马、狗、鹰、蝙蝠、猪、鳄鱼以及其他很多动物的前后肢骨骼基本相同。在进化中，仅仅改变骨骼的尺寸，很多特殊的功能就可以成为可能，例如，相对修长的四肢骨是马快速奔跑的秘诀，而相对较轻的翅骨则有利于鸟类的飞行。

四肢的存在与某个在许多生物体内负责塑造形态的调节因子家族有关，物种跨度从水母到人类。虽然这些调节因子是生物躯体正常发育所必需的，但是编码这些调节因子的基因，即homeobox或Hox，它们的名字是根据它们在同源异形现象（homeosis）中的作用而定的。同源异形现象指这些基因变异后造成的生物畸形，例如畸变的果蝇在头部原本是触角的位置长出了无用的腿。通常来说，改变生物正常的基因表达往往会引起严重的后果。

homeobox是一种含有60个氨基酸的蛋白质，它能够与DNA结合从而帮助Hox调节因子调控基因的表达。无论是在果蝇还是人类体内，这些调节因子都主宰着其他数百种与细胞、组织以及器官形态有关的基因活动。除了单个基因之外，Hox调节因子还在调控另一种东西，那就是调控环路。如果你还记得图5-4的话，Hox家族的调节因子参与的调控环路要比图5-4里的例子复杂得多，因为动物细胞中的调控环路通常包含40种乃至更多种调节基因。类似的环路往往是决定动物身体形态的关键，其中也包括人类。人类脊柱中的33块脊椎骨以及它们独特的形态，即脊椎颈部的前两块骨头构成一个灵活的关节，12块胸椎与肋骨结合处的关节槽等，就是典型的例子。

胎儿在子宫内发育的时候，Hox基因家族就已经在调控颈椎、胸椎和腰椎的基因表达了。每一个部位的Hox家族各个基因的开与关就构成了这个部位的基因表达代码（gene expression code），不同部位的表达代码各不相同，某个表达代码代表颈椎，另一个则代表胸椎，诸如此类。

Hox基因不仅负责塑造人类的身体，它也在其他脊椎动物，例如蟒蛇或任何种类的蛇体内参与形态构建。蛇的独特形态是大自然无心插下的又一根柳条，使得这种生物能够在地面蠕动、在地底穿梭、在水里游

弋。某些种类的蛇拥有超过300块脊椎骨，它们中大多数的躯体结构与我们的12根胸椎无异，也连接着肋骨。Hox基因家族便是蛇与其他脊椎动物产生形体差别的原因：在大多数脊椎动物体内，Hox家族只在胚胎的一小块区域内激活胸椎基因的表达，但是当蛇的进化之路在大约1亿年前与蜥蜴分开之后，这一小块区域就像拉开的橡皮筋一样扩展开去。指导胸椎形成的Hox表达出现在了身体中轴上的大部分区域，使得这些区域的数百个脊椎发育为胸椎，继而造就了蛇独特的身材。

Hox基因家族不仅在动物中轴骨的形成中起着决定作用——在脊椎动物里，这里所说的中轴骨基本相当于脊椎骨，它们还参与了另一个脊椎动物的结构形成：鱼鳍。鱼鳍并不是一成不变的。在过去的数百万年里，进化通过Hox基因家族的变异、优化和分化，逐渐把鱼鳍变成了四肢。不管生物是在地上奔跑还是在天空翱翔，四肢骨的基因表达都可以被划分为相似的三部分，第一部分的Hox基因控制上臂的形成，第二部分控制前臂的形成，而第三部分则控制手掌的形成。

某些基因的变异在动物体内导致了可怕的先天缺陷，上述的结论正是来自对这些变异缺陷的研究。比如发育的四肢中如果缺乏Hoxa11和Hoxd11的表达，那么结果将导致胚胎没有前臂，手掌从手肘的位置萌发出来。同样的道理，如果Hoxa13和Hoxd13表达缺失，那么胚胎就可能无法萌发手指或手掌。还有一些Hox基因表达的缺失则会导致前臂和手掌同时消失，仅有上臂能够正常发育。

当然，正常情况下Hox家族的基因都还是恪尽职守的。此外，它们所参与控制的结构数量惊人，从盆腔到颅骨，都离不开Hox家族的身影。Hox基因家庭同样存在于虾、水母、蠕虫甚至是果蝇体内，而且重要程度几乎与控制体节形成的调节因子家族相当。实际上，两者的工作是前后衔接的。当体节家族的成员敲定体节的数量之后，Hox家族继而决定每个体节的功能，如这个体节负责腿，那个体节负责翅膀等。体节和Hox基因家族还只是果蝇和其他动物众多同类型分子家族中的两个，从数亿年前动物出现伊始，这些家族就已经参与到对形态的控制中了。

Hox基因家族对于动物新形态的诞生至关重要，例如从鱼鳍到四肢的改变；还有新的中轴骨，比如蛇的身体。关于这些新的形态究竟是如何起源的，也许已经随着生命漫长的历史，永远故去在风里了，但是仍然有一条原则亘古未变：新形态的起源必然伴随调节方式的改变。

不仅是形态，这个原则在所有新性状的起源里都应该适用。

这里我们可以想象有一条纤细的蜥蜴，它在茂盛的草丛里蜿蜒游走，寻找下一顿美餐。突然它呆住了，因为它发现面前出现了一双巨大的眼睛正死死地盯着它。它意识到自己可能马上就会被眼前的怪物撕成碎片。但在突然之间，面前的眼睛像海市蜃楼一样消失了，只见一对扇动的翅膀乘风而去。原来根本不是什么捕食者，而是一只美丽的蝴蝶和它翅膀上的两个巨大的色斑而已。

蝴蝶的眼状斑点是它的保命伎俩，由一种非同寻常的调节蛋白distalless控制。作为调节环路中的一员，除了参与果蝇的腿、翅膀和触角的形成之外，distalless还给蝴蝶的翅膀画上了奇异的眼斑。我们之所以能够确定distalless和独特的眼状斑点之间的关系，是因为distalless在蝴蝶幼虫发育中合成的位置正好与将来眼斑形成的部位吻合。有的蝴蝶的眼斑十分巨大，有的则相对较小，有的只有一个眼斑，而有的则有数个。但不管是哪一种，发育中的蝴蝶都在眼斑的位置上持续表达着distalless。distalless合成与眼斑位置的吻合不是巧合，实际上它的确是眼斑形成的原因：如果把合成distalless的细胞移植到翅膀上的其他位置，最后眼斑也会出现在同样的位置上。

蝴蝶的身体就像一座大教堂，从体节的中殿到眼斑的滴水嘴兽，都是基因调节一手造就的。基因调节是优秀的建筑大师，不管图纸有多复杂都不成问题，其中也包括植物以及它们的根、茎、花、叶。第一株开花植物出现在大约两亿年前，它们的叶子边缘齐整，叶面平滑连续。随着时间的推移，单叶逐渐进化成深裂叶，一张叶片分裂为许多小的叶片（如图5-5）。

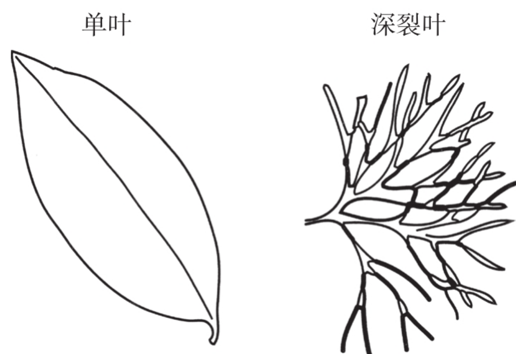


图5-5　叶形

单叶进化为深裂叶给植物带来了许多优势。深裂叶的表面积比单叶要大，这使得叶片能够吸收更多的二氧化碳进行光合作用，从而促进植

物以更快的速度生长。此外，更大的散热面积能够使叶片在炎热的环境里不至于过热，而过高的温度往往会抑制光合作用的速度甚至损伤叶片。如果深裂叶这么好用，我们猜测它可能在进化上起源过不止一次。事实确实如此，仅仅在开花植物的历史上，深裂叶就有过至少20次相互独立的起源。

每一次深裂叶的起源都伴随着基因调节的变化。植物在萌发的初期，只有尖部一小块组织内的细胞能够分裂，促进幼苗生长，推动植物向上穿破土壤。这块分裂组织也是一株植物所有的叶子最初起源的地方。在形成肉眼可见的幼叶之前，一小簇细胞，或者叫叶原基，就已经从其他分裂的细胞中分化出来，专门用以形成叶片。叶原基里的细胞都会表达一种名为KNOX的调节蛋白。

牛津大学的安杰拉·海（Angela Hay）和米托斯·茜提斯（Miltos Tsiantis）曾经用一种十分不起眼的草本植物——碎米荠（hairy bittercress）进行实验，通过控制植物体内KNOX蛋白的合成，他们才发现了这种蛋白质在叶片形成中扮演的重要角色。随着KNOX的合成量降低，碎米荠深裂叶的缺刻数量会逐渐减少直至变成单叶。而如果调高KNOX的合成，深裂叶的缺刻就会相应增加。不仅如此，他们还发现KNOX不仅在碎米荠中，还在许多其他种类的植物体内控制着深裂叶的形成。



上面的那些例子，以及无数我们未提到的事实都反映了基因调节对于生物进化的重要性。无论是世界上各个实验室的研究记录本，还是各种学术刊物，都充斥着有关基因调节分子的研究，比如植物中的KNOX、蝴蝶体内的distalless以及果蝇的齿状蛋白等。人类的基因组编码了超过2 000种不同的调节因子，它们构成了数十个相对独立的调控环路。过去半个世纪的研究已经让我们窥见了基因调控在塑造生物形态中的重要性，它有助于我们理解许多新性状进化的过程，以及性状背后的基因表达代码。

但不论现实中的例子有多丰富，也不过只是告诉了我们蜥蜴的四肢和鱼的鳍在发育中与Hox基因家族有关而已，即不同的调节因子表达谱导致了不同的基因表达结果。即使我们找到了新性状与新调控环路之间的关联，也还是无法解释进化是如何找到这些最合适的基因表达的。

（调控环路的种类越多，要弄清这些环路的起源就越困难。）另外，由于调控环路在进化过程中时刻积累着微小的变化，如何保持已有的优良基因表达谱就成了一个充满矛盾性的难题。仅有调控促进进化的例子，还是无法告诉我们这个过程到底是如何实现的。

你可能会觉得这个问题很眼熟，其实它的答案也不陌生：我们需要研究尽可能多的调控环路，最好是整个图书馆里的调控基因型和它们的表现型。调节因子图书馆里收录的是编码调节因子的DNA，以及它们识别的DNA关键词。但如果我们直接以这种方式记录所有的馆藏，整个过程将无比烦琐和冗长，就像你要用每一个分子的空间定位来描述一栋房子一样。其实你大可以用一张房子的图纸省下很多力气，就像图5-4中的示意图那样。

整个调节图书馆内包含了所有可能的调控通路，换句话说，就是包含了所有可能的图纸。你可能会觉得归档这些内容很困难，但实际上极其简单。调控环路里的任何一个调节因子，比如说A，如果我们要研究它与另一个调节因子B的关系，那么它的影响不过就是3种可能中的一种：A能够激活B、抑制B，或者根本对B没有影响。对于任何一对调节因子而言都是同样的道理，比如图5-4中的调节因子A和C，或者D和E，调节因子能够激活、抑制另一个调节因子，或者毫无作用。

对于所有调节因子来说都仅有3种可能性。这个基本的原则能够帮助我们更好地理解图5-4中的5个基因的调控环路，接下去要做的就是数清这5个基因之间有多少种配对方式。图5-4中的环路内有5^2种配对方式，对于每一对基因而言都有3种不同的可能效应。第一对基因有3种可能的效应，第二对、第三对也同理，以此类推，直到第25对。所以5个基因的调控环路的所有可能结果一共有3^{25}种，换句话说，由5个基因组成的调控环路有超过8 000亿种可能。

对于5个基因来说，8 000亿这个数字显得着实惊人，尤其是许多现实的调控环路中包含的基因数量远远不止5个。以脊椎动物的Hox基因家族为例，它们组成的调控环路里有至少40个基因。要计算40个基因的调控环路有多少种可能性，我们可以采用同样的方法：首先计算基因的配对数量，为40^2（1 600）种，然后计算$3^{1\ 600}$。如此大的量级对于我们而言并不陌生，这个数值超过10^{700}，如果把它印刷出来，那么可以铺满这整页纸。

但请不要忘记的是，虽然这个数字已经超乎常理的大了，但是它和环路数量的实际值相比依旧有差距。因为到目前为止，我们一直假定每

个调节因子在调控环路里的作用是同等重要的，一个调节因子要么将目标基因激活，要么将其关闭。事实上，有的调节因子的作用相对较弱，有的则相对较强，作用强弱的差别使得情况大大复杂化了：每对基因面对的可能结果不再是3种，而是5种：没有作用、弱激活、强激活、弱抑制或强抑制。于是，我们计算的幂底数就从3变成了5。这还没有完。

如果我们有办法进一步区分基因激活或抑制的强弱程度，那么可能的调控环路数量还会继续增加。幸运的是，我所在实验室的研究表明，对于激活或者抑制程度的细分除了数量之外，并不会改变整个图书馆里的组织原则。这是个好消息，说明数量根本没有那么重要，因为光是以3作为幂底数，调控环路的数量就已经是超宇宙级别了，再多一些似乎也无妨。

基因调控环路图书馆和它收录的基因型馆藏与我们之前探讨过的代谢图书馆和蛋白质图书馆有诸多相似之处。当基因发生变异之后，我们以添加或是去掉基因之间的线条来表示两者调节关系上的改变。但是请记住，这些线条不是真实存在的，仅仅代表基因之间存在调节关系，而这种关系受到变异的影响。每当你改变其中一对基因的调节关系，你就得到了一个原环路的相邻环路。

图5-6中给出了一个例子，由于变异，与左侧的环路相比，右侧环路中的B基因不再调节D基因的表达（这种调节在左侧由那个粗箭头表示）。每个调控环路都有许多相邻环路，如果是40个基因的调控环路，每种环路的相邻环路将达到3 000多个。如果我们把所有的调控通路安置在一个超立方体上，每个顶点对应一个环路，再故伎重演，那么寻找相邻环路的过程相当于沿着超立方体的边从一个顶点移动到下一个顶点。由于超立方体存在于多维空间，所以从每一个顶点出发的边有许多条，与40个基因对应的超立方体存在于1 600维空间。这个超立方体的顶点数量远远不止1 600个，而是达到了10^{700}个，这也是图书馆中所有包含40个基因的调控环路的馆藏数量。

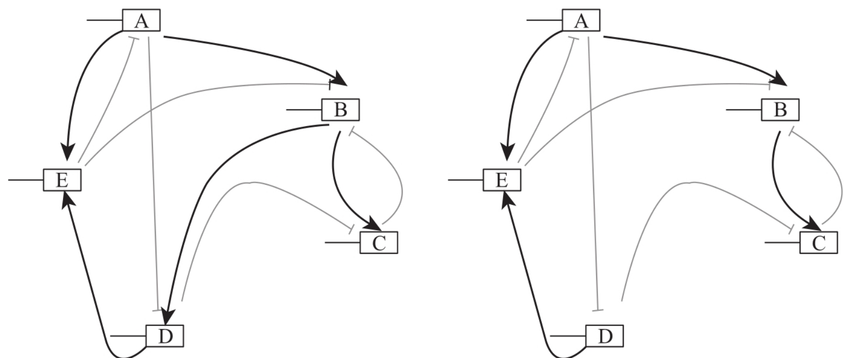


图5-6　调节环路图书馆中的相邻环路

和我们前面介绍过的其他两座图书馆一样，超立方体上的每一个环路都有各自的“社区”，里面包含所有与之在图书馆里相邻的馆藏，也就是仅相差一对或少数几对的调控环路。微小的基因改变，哪怕只是一个DNA分子的变异，就有可能建立或是摧毁一对基因之间已有的作用，进化只要抬脚走上没几步，就能从一个书架走到下一个书架并浏览上面的文本。循着相邻基因一直往前，你就会逐渐深入到图书馆内——这样的旅程你已经不陌生了。而在这里，距离的概念变成了两个环路之间连线的差别。越是临近的环路之间差距越小，而相距甚远的环路之间则几乎没有相同的连线，分别位于图书馆中截然相反的两个方向。

同样的道理，图书馆里的多数环路基因型是随机的，没有任何意义。但也有一些编码了有意义的单词或句子，只是整体而言依旧词不达意、不知所云，甚至会宣扬恶俗言论，比如变异的Hox基因最终将导致没有手掌的残疾手臂。我们这里所说的文字和语言，同样是指基因调节和表达的化学语言，只有细胞和组织真正理解它们，并最终将它们翻译成脊椎骨、叶子或手掌等血肉的语言。而新性状的诞生过程我们在单叶进化为深裂叶中已经多少介绍过了。

我们在前面已经探讨了调控环路通过操纵基因表达控制性状的原理。从一套预先存在的调节因子开始，比如果蝇在受精卵内留下的分子信号，后续的调节因子逐渐形成调控环路，并改变最初的基因表达模式。基因在发育过程中开开闭闭，直到抵达某个平衡点，然后就犹如马戏团里的杂技演员们，保持巍然不动。对于马戏团里表演叠罗汉的杂技演员们来说，他们的平衡建立在相互牵制的基础上，这一处的推力在那边相当于拉力，而打破这种平衡状态唯一的办法则是瓦解其中的某个个体。

通过多年的研究，我们对于这种平衡的理解已经足够我们演算平衡点了，就像约翰·瑞尼茨的果蝇模拟器。我们已经能够同时考量的环路数目不是一个或几个，而是数百万个，这相当于同时演算整个超宇宙级别的图书馆。

我们从一开始就知道环路图书馆里的馆藏数量超出任何人的想象，哪怕是环路的表达谱数量也不是闹着玩的。如果在一个含有40个基因的环路中，每一个基因只有激活或关闭两种可能，那么40个基因就有2^{40}种可能表现型，总值超过一万亿。而现实中一个基因的状态并不是非白即黑的，它可以微弱、中等、强烈或是非常强烈地进行表达。不仅如此，生物形态的造就往往需要多个不同的调控环路协同合作，这也大大增加了表达谱的可能数量。与所有这些表达谱的数量相比，我们体内区区数百种不同的细胞和组织几乎不值一提。如果我们把体内所有的细胞都铺陈出来，让每一个细胞对应基因表达谱中的一种，那么最终将无法容下所有的表达模式。

进化用我们熟悉的随机游走方式探索着调控环路的图书馆，生物种群的形态重塑来自偶然的**DNA**复制错误，这些错误的复制一般发生在亲本将遗传物质传递给后代的时候。微小的突变通常会导致两种可能的结果，即改变调节因子的形态并阻止它们与DNA结合，或者直接改变DNA上调节因子可识别的“关键词”——这种改变会阻碍正常的调节因子识别对应的基因，抑制基因表达，同时也有可能令DNA被新的调节因子识别。

上述的第一种结果往往会造成灾难性的后果，因为一种调节因子通常可以作用于许多种不同的基因。如果调节因子失去识别DNA的能力，相当于把一份食谱里的原料混淆一气，最后做出的料理可想而知，这会导致生物体的严重畸形，甚至胚胎在出生前就会夭折。而第二种结果则更像是食谱里的某个印刷错误，往往只涉及某个基因的表达以及相应的蛋白质数量——它只不过是数千种蛋白质中的一种而已，这使它导致严重后果的可能性变得很小。有人可能会想，生物体对第二种变化的容忍度要更高，因此也更容易在进化的时间跨度上稳定地积累下来。如果当真如此，这些积累的微小变化就能够逐步改变环路里的调控模式。

如果把在过去数百万年中独立进行进化的调控通路拿来比较的话，比如数千种不同的果蝇体内的某几个调控环路，我们就会发现，生物体耐受性最好的变化发生在环路内的基因之间的相互作用中，而不是基因本身。进化的改变总是从某一对基因之间的作用着手，因为直接对基因

下手容易造成严重的后果。此外，基因对之间作用的微小影响的确会积累并逐渐改变调控环路，而这个过程十分漫长。改变缓慢的原因在于调节因子的DNA关键词通常只有5个碱基对的长度，且与下游基因有着数千个碱基对的距离。如果仅凭概率，那么随机突变产生新关键词并由此将两个基因联系在一起的可能性要更大一些。



如果调节环路图书馆中的10^{700}件馆藏里只有一个表达谱与Hox家族吻合，那么它就如同一根掉进宇宙的绣花针，生物进化大可以早早放弃挣扎。我在20世纪90年代就疑惑过，为何进化最终还是战胜这渺茫的概率找到了Hox家族，不过我并没有把这个问题太当回事，当时的我正为别的研究项目忙得焦头烂额。直到2004年，当我在靠近法国巴黎的高等科学研究所进修休假[17]时，才开始认真考虑这个问题。

高等科学研究所坐落在一个到处是参天古树的田园乡间，那里有着精心修剪的灌木、争相怒放的花朵，还有在思考问题时能够信步的幽静小道，简直是对那些受尽经费申请、社交应酬以及社区活动纷扰的生物学家来说最好的避难所。常驻在研究所内的几个为数不多的科学家都是非同凡响的优秀学者，他们当中有数名菲尔兹奖的获得者（菲尔兹奖是公认的数学领域的诺贝尔奖）。

研究所主要的研究领域是数学和物理学，但所内的首席科学家们早就注意到了分子生物领域长期以来的停滞不前，他们预见到当前盲人摸象般的研究方式需要一种整体水平的学科进行整合，而这种洞见的产物就是一门分支学科：系统生物学（systems biology）。这个新近出现的研究领域把实验数据和数学、计算机技术结合起来，试图解释分子水平的活动如何作用于生物整体，换句话说，就是微观分子如何构成了宏观生物体。数学家和物理学家手握许多解决这种问题的理论手段，所以研究所邀请了像我这样的生物学家造访，想要看看不同领域的科学家能否联手合作。

我很庆幸当初接受了这份邀请，因为正是在巴黎，我遇见了奥利弗·马丁（Olivier Martin）。

奥利弗是一名享有国际声誉的统计物理学教授，任职于巴黎市郊的奥赛大学（University of Orsay）。像奥利弗这样的统计物理学家，很擅长解决海量微观粒子的宏观现象，比如高压密封罐里的丙烷气体是如

何在宏观上产生压力的。对类似压力现象的把控非常重要，谁都不会愿意看到储气罐爆炸，但想要实现也并不容易，因为气体分子和储气罐内壁每时每刻都在发生万亿次的碰撞。统计物理学家乐于把数以万亿计的微观分子看作一个整体，因为单独考量每个个体分子几乎是不可能的，所以他们发明了一套用于研究类似体系的统计方式，算法中包含了复杂的统计学手段。不过我们这里所说的统计分析除了名字之外，与美国大选中民调专家口中所谓的统计分析没有任何关系。

奥利弗也有自己的烦恼。今天的统计物理学就像一家被饿汉们席卷之后的自助餐厅：大多数主要的问题基本都已经被解决了，剩下的都是残羹冷炙，剩下的问题不是太艰涩就是太微不足道，这样的情况从19世纪詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）和路德维希·玻尔兹曼（Ludwig Boltzmann）用统计学开创热力学以来就没有发生过太大的改变。和他所在领域里的众多其他科学家一样，奥利弗非常渴望突破物理学的局限。他的烦恼在于不知道如何在系统生物学领域找到一个难度合适的新问题，让他所掌握的统计物理学技能找到用武之地。

而我却有一座馆藏数量达到10^{700}件的图书馆要照看。乖乖，我就是奥利弗的贵人啊！

奥利弗·马丁和我开始合作之后没多久，我就对他感激不尽。他作为科学家的直觉和学术素是为我们探索图书馆最好的保障。不过他可不仅仅是一位旅途中的好伙伴，更是一位亲切慷慨的师长，不时耐心地提点我们，用他掌握的专业技能帮我们摆脱眼前的困境。

我们研究开始的第一步是为了解决一个问题，而这个问题你已经非常熟悉了：某种含义的文本在调控环路图书馆里是不是唯一的？为了找出答案，我们从图书馆里的某个环路开始，模拟和计算这个环路的表达结果。然后我们改变其中一对基因的作用关系，查看这个变异能否改变表现型，接着再恢复到最初的环路，改变第二对基因，依次类推，直到我们检验完所有的相邻环路并得到它们的表现型。为了排除某个调控环路的偶然性，我们选择了许多不同的环路作为起点进行上述的检索，这些环路包含了不同的基因数量、不同的相互作用基因对、不同的基因相互作用以及不同的表现型。

最终，我们得到的结论是一致的。每个调控环路所处的社区里往往包含数十个到数百个表现型相同的相邻环路。也就是说，即使发生基因突变，单个基因对作用关系的改变也不一定会引起调控环路表现型的改变。基因调控环路不像马戏团里杂技演员表演的叠罗汉，不是其中一个

个体“差之毫厘”，整体就会“失之千里”。调控环路的基因型之所以能够耐受这样的突变，是由于并非每一对基因间的关系都对环路的功能有重要贡献。

从一个环路出发的第一步我们就已经得到了一个重要的结论：没有哪种性状，不管是果蝇的体节、植物的深裂叶还是脊椎动物的脊柱，都是由独一无二的调控环路塑造的。基因对关系不同的调控环路同样能够产生相同的性状。不过我们不知道这些表现型相同的环路到底有多少，在超过40个基因的调控环路里筛查所有环路的计算量是我们力所不及的。哪怕是规模小得多的环路，计算量也非同小可：10个基因的调控环路有超过10^{40}种可能性，而20个基因的环路则有10^{160}种。实现同一个性状的基因型远远不止一个。

接下来，为了评估同一表现型的不同基因型在图书馆里的相对距离，我们再一次借助探索代谢图书馆和蛋白质图书馆时用过的随机游走。选择一个起点，模拟环路的表现型，而后改变某一对基因的作用，即添加或者抹掉一个基因对另一个基因的作用，以此移动到相邻的调控环路上，在保证表现型不变的情况下，继续重复上述步骤，直到无法通过改变任何单个基因对维持表现型稳定。

不出所料，我们又能凭借这种方式横穿几乎整个图书馆。基因对差异超过90%的两个调控环路依旧能够产生相同的表现型。如果用示意图标记两个环路里的基因联系，你可能永远想不到两者是如何经过一步一步微小的变化而成为对方的。但它们的确指向了某个问题的答案：如何产生某种特定的蛋白质指导细胞特定的分化。

为了确保结果具有代表性，包括我们模拟的表现型，我们尝试了以各种不同的环路作为起点，不同的环路基因数量、不同的基因间关系数量、不同的作用关系以及不同的基因表现型，而这些对结论都没有影响。某些环路虽然差异巨大，但是表现型却相同，最小的差异只有“区区”75%而已。但是“如此小”的差异依旧很难让人联想到它们之间居然还存在着联系。

进化还告诉我们，所有图书馆中表现型相同的调控环路是相互联系的。我们可以以它们中的任意一个作为起点，通过一次改变一个基因对并保证性状相同的方式，检索到表现型相同的任何其他环路。和在代谢图书馆以及蛋白质图书馆里一样，我们在环路图书馆里又找到了一种从一点走到几乎任意一点的方式，而不至于迷失和身陷在无意义的环路泥潭里。

于是，调控环路图书馆里所有表现型相同的调控环路也形成了一张巨大的网络，类似的基因型网络我们已经在代谢图书馆和蛋白质图书馆里接触过了。调控环路图书馆里充斥着这样的网络，每一个网络里都包含了数不清的环路，零散分布在图书馆各个角落。同一张网络里的每个调控环路都有相同的作用：指导特定的基因表达谱，帮助特定的细胞、组织或器官分化。一张网络中只有包含足够数量的调控环路，不起眼的深裂叶新性状才有可能在进化史上不断重复上演。

如果没有计算机的帮助，通过对数百万个环路的模拟来理解环路图书馆几乎是不可能的，这意味着数百名实验科学家必须要花费数十年，在数百万种果蝇身上完成实验，才能理解一种调节果蝇体节分化的基因环路。但是，的确有一些胆识过人的科学家在某些更低等的物种身上进行这种研究工作，他们的研究对象是细菌和真菌。巴塞罗那科学家马克·艾莎兰（Mark Isalan）就是其中之一，他在大肠杆菌的调节因子基因间构建新的联系，创造了数百种大肠杆菌环路的相邻调控环路。和我们的演算结果一样，他发现调控环路对于内部基因的关系变化具有相当的耐受性。他在大肠杆菌身上构建的改变中有95%没有引起任何功能性变化。

还有科学家比较了多种啤酒酵母体内的调控环路，想知道它们之间的差异究竟能有多大。这种调节环路的其中一个作用是激活消化乳糖的乳糖酶基因。你可能会想，类似的调控环路应当具有某些共同特征，只有找到发现和运用这种特征的真菌才能拥有相应的代谢能力，并把这种能力稳定地遗传给后代。而事实并非如此。以两种进化道路在数百万年前就各自独立的真菌为例，不仅它们的调控环路完全不同，甚至于参与环路的调节因子都不一样。两种调控环路不分高低，如若不然，它们也无法同时保留至今。自然界以两种不同但是同样有效的方式解决了同一个问题。不仅如此，用于功能相同或不同的调控环路之间，也被一步步微小的变异联系在了一起。

核糖体能够将RNA翻译为蛋白质，编码这种分子机器的基因向我们诉说了同样的故事。细胞在高速合成蛋白质的过程中，必须精确控制不同蛋白质间的数量平衡，不然就可能会像过量合成无用β-半乳糖甘酶的大肠杆菌那样破裂死亡。维持这种精妙的平衡似乎很难，很容易让人误以为只有某种最佳的解决方式才能实现。但是实际上，有两种不同的真菌分别以两种完全不同的调控方式实现了这种平衡。

类似的例子向我们展示了生物对调控环路图书馆的深入探索。但是

物以稀为贵，在寻找新颖和高效的表达谱的过程中，生命面临着它们在探索代谢图书馆和蛋白质图书馆时就曾遇到过的同样的难题：环路图书馆里有数万亿种不同的调节环路，但与某个环路直接相邻的社区中却最多只有几千种环路，以这种方式寻找新的基因型效率着实低下。

为了获得尽可能多的新性状，调控环路的进化必须设法跨出所在的社区。这种探索图书馆的方式只有在不同的社区间存在巨大差异的情况下才能收益最大化。为了验证事实是否如此，我们让计算机从一张基因型网络中随机挑选两个调控环路，姑且称之为A和B，它们的表现型相同，但是环路结构天差地别。接着，我们再找出它们各自所有的相邻环路，并分别演算它们的基因编码。我们发现，在A环路与B环路所在的社区中，大多数相邻环路所控制的基因表达谱都不相同，更不要提A环路和B环路本身在基因数量和基因关联上的巨大差异。不同社区中的环路往往表现型也不同。

于是我们的故事又回到了熟悉的套路上。环路图书馆的布局与代谢图书馆和蛋白质图书馆类似。我们把指导相同基因表达谱的调控环路安置到一张巨大的基因型网络里，对于在这张网络里漫无目的游荡的读者而言，他们只能象征性地在“换汤不换药”的馆藏里寻找新书。真正推动生物进化的动力恰恰是无头苍蝇一般的随机突变，调控环路在稳定积累的微小变异中逐渐改变：虽然某些突变葬送了整个调控环路，但依然有一些突变在保留环路功能的基础上把生物推向了不同的基因型社区，获得了新的基因表达谱，而总有一个表达谱会为下一步生物形态的重塑埋下伏笔。我们再次看到，基因型网络中多样的基因社区成为新性状出现的关键。

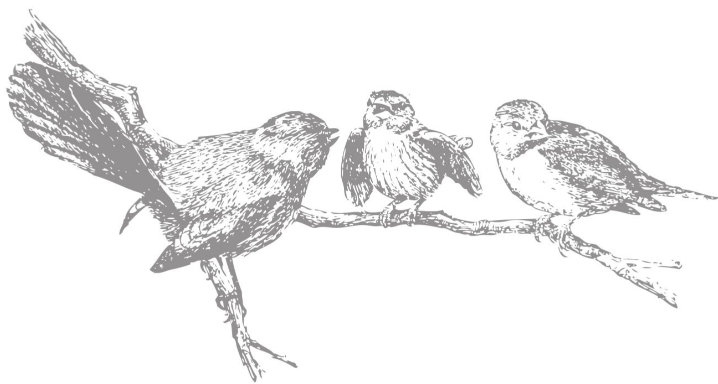
不同图书馆之间的这些共同点让人捉摸不透。为什么代谢、蛋白质和调控环路图书馆中的新性状有着相似的起源方式？为什么不同的图书馆却有着十分类似的分类体系？这些问题的答案是一只看不见的手，它早在生命出现之前就在引导着世界万物的进程。这只手是自组织，而生命的自组织作用显得尤为奇特，接下来我们就回到这种作用上。

ARRIVAL OF THE FITTEST

06 神秘的建筑师

Solving Evolution's Greatest Puzzle

多变的环境催生了生物的复杂性，而复杂性促成了发育稳态，发育稳态继而造就了基因型网络，后者让进化成为可能，使得生物能够通过演变适应环境的变化、提高自身的复杂性，循环往复，生物进化通过这种方式螺旋上升。这种进化方式的核心在于处在多维空间的基因型网络的自组织性。自组织性是生命绚烂光彩背后的支持者，它是隐藏的生命建筑师。



1944年，诺贝尔物理学奖得主、理论物理学家埃尔温·薛定谔（Erwin Schrödinger）出版了《生命是什么》（*What is life*），书中收录了他所做的一系列演讲。在那个沃森和克里克还没有发现DNA双螺旋模型的年代，薛定谔已经开始尝试从物理学的角度解释所谓的生物进化。《生命是什么》一书虽然篇幅不长，但是充满了真知灼见。这些洞见中有一个一直受到主流科学界的广泛认可：进化增加了有序性，减少了无序性，或者用薛定谔自己的话说，叫“负熵”（negative entropy）。

4年后，美国电气工程师克劳德·香农（Claude Shannon）把本来只在热力学领域使用的概念“熵”（entropy），借用到了电报的信息传输精度问题中。于是，进化和信息之间的关系就被建立起来了，只不过这种关系的表述非常简单粗糙：无序——不好，有序——好；正熵——不好，负熵——这时候也叫信息——好。

自从薛定谔出版了《生命是什么》，我们对熵的理解就开始变得更加复杂而严谨。有序性和信息传递一直都是生物进化的中心议题，但是近几年来，基于对基因型网络的认识，我们发现，完美的有序性和彻底的无序性一样，对生物进化而言都有害无益。无序性对自然来说并不一定意味着负担，它同样可以帮助自然界的生物发现新的代谢方式、新的调控环路和新的生物大分子。简而言之，无序性也可以帮助生物进化。

我们再拿乐高积木来打个比方。乐高积木可以根据孩子的想象力随意拼接，当然，这些我们熟悉的塑料块也可以根据预先设计好的“图纸”拼出特定的形象。例如乐高公司可能会给孩子们提供一张图纸，只要他们按照图纸上的步骤就能拼出一艘海盗船。无序的拼接比照搬图纸更有可能创造出新的形象，这种潜力一方面要归功于孩子们丰富的想象力，而另一方面则是因为乐高积木有许多可能的海盗船拼法，远远不止说明书上列出的那些。

同样的原理在生物学上表现为自然界的基因型网络，即相同表现型的不同基因型集合，比如鳕鱼体内的抗冻蛋白。在更隐晦的层面上，基因型网络的存在牵涉到一个广泛存在的生物学概念，但是这个概念曾长期被人忽视，直到20世纪末才引起科学家的注意，这个概念就是发育稳态（robustness）。发育稳态指面对外界环境时生物体保持自身特征稳定的能力。

要理解生物学中发育稳态的真正含义，最好的例子莫过于把传统印

刷出版物和计算机程序中的输入错误进行比较。如果在一本书里看到这么一段文字：

N smll stp fr mn, n gnt lp fr mnknd[18]

读者可能会眉毛一扬，然后若无其事地继续往下读，因为对于一本书而言，这么一小段文字乱码几乎不会影响对整本书的理解和阅读体验。然而在计算机程序里就不一样了，对于动辄几千行的代码来说，不要说字符，哪怕是一个逗号的丢失都有可能让价值数百万美元的程序崩溃。在现实生活中，类似的程序错误每年都要造成数十亿美元的经济损失。相较而言，我们可以说，人类的语言具有很强的发育稳态，而计算机的程序语言则没有。

生物具有发育稳态的猜测最早可以追溯到20世纪40年代，当时的生物学家兼哲学家沃丁顿（C. H. Waddington）在研究果蝇时发现，不同基因型的果蝇在身体外观上几乎没有区别。沃丁顿用于观测比较的指标非常精细，例如果蝇翅膀上的脉络以及背部的刚毛数目。沃丁顿根据这种现象认为，“发育无论在什么情况下都会止于一个相同的最终结果”，并将其定义为限渠道化（canalization）——这个词表达的含义与发育稳态相同。沃丁顿的研究暗示，果蝇的形态对基因突变具有一定的耐受性，即同一个身体性状背后的基因型可以有很多种。即便如此，对于发育稳态的研究在此后的半个世纪里仍旧是一潭无人问津的死水。

到了20世纪90年代，当时的分子生物学家们深受一个问题的困扰：他们发现许多基因根本没有任何实际作用。这个表面上看起来与沃丁顿的研究没有任何关系的现象，让发育稳态这个概念几乎在一夜之间就涌入了主流科学界的视野。

科学家们想不明白的是，既然这些基因没有用，那么它们为什么没有消失呢？这些无用基因的存在浪费了宝贵的资源，不断积累的DNA突变应当将它们抹去。就像一栋被人弃用的大楼，年深日久，最后必定会归于尘土。

我们在第5章中提到过啤酒酵母，除了用来酿酒，它对细胞生物学的重要性犹如果蝇之于胚胎学。正是在对这种啤酒酵母完成全基因组测序后，科学家们才发现了众多的“无用基因”。手握酵母的整个基因组，科学家们意识到，他们对基因组中数千个基因在酵母生命中扮演的角色一无所知。为了了解每个基因的作用，科学家们对酵母的基因进行了逐个“敲除”（knockout），之所以这么命名，是因为科学家每次只删除基

因组里的一个基因，相当于从基因组的书里删除某个表意完整的特定段落。

这个研究的逻辑就像通过每次去掉汽车的某一部件，从而分析该部件在汽车里的作用一样：如果你拆掉车的圆盘转子再踩刹车踏板，发现车不能减速，那么你就能猜到转子和汽车的刹车有关系。同样的道理，如果敲除某个基因后酵母不再分裂，那么这个基因就和细胞的分裂有关。而如果被敲除某个基因后的果蝇失去了翅膀，那么这个基因就参与了果蝇翅膀的发育。

随着一个又一个有关基因的科学论文发表，基因敲除技术发展到今天，已经强大到能够敲除数千个不同的基因。斯坦福大学的研究者们在20世纪90年代首创了这种惊人的研究方法，他们在获得酵母基因组所有的基因测序结果后，决定逐个敲除它们。研究人员创造了6 000多种不同的酵母变种，然后把它们培养在未突变菌株能够生存的环境里，检测所有变种存在的特定缺陷，以此推断敲除基因的可能功能。

研究人员的发现出乎所有人意料。大多数变种在培养基里没有表现出明显的缺陷，它们与非变种一样生长旺盛。这也就是说，大多数敲除的基因没有什么实际作用。以此为开端，科学家在许多其他物种中干预了无数基因的表达。而得到的结论是，基因就像一句不含一个韵母就写成的句子：生命如同人类的语言，有很强的发育稳态，因此能够耐受实验中多数的基因敲除。

这样的结论引发了更多的疑问，其中之一便是，为什么会这样？发育稳态背后的原理是什么？

对于某些基因来说，发育稳态的原理简单明了。在基因组里，一些基因往往存在多个拷贝，就像有人在影印书籍的时候不小心把某一页重复印了两次。基因重复通常发生在DNA复制和修复的过程中，而且并不少见：人类基因组中大约有一半的基因都存在重复。由于重复的基因有相同的作用，所以当其中一个基因被敲除时，其他的拷贝就能够补上空缺。就像医院里防备停电的备用电源、用于备份数据的计算机硬盘或者商业航班中防止坠机的备用电路一样，如果不需要这些基因补上空缺，那么它们就一直都是“无用”的。

但鉴于多数“无用基因”没有复制——它们是单拷贝基因。因此，对这些基因而言，发育稳态现象并没有上述那么容易解释。

对于多数单拷贝基因，我们知道有一种情况可以解释它们的无用

性，而这种情况普遍存在于催化代谢的酶中。生物体内的生化反应网络有点儿类似于城市中心的密集交通网络，比如曼哈顿市中心纵横交错的公路和街道。一名位于第二大道42号街的司机如果想去第七大道的48号街，可以选择北边东西走向的6条街道以及西边南北走向的7条街道中的任意组合。城市里的每条主路通常都有数条车道，车道越多，备选的前进路线就越多，就算有一条车道堵死，司机也可以选择走其他车道。不过即便整条路都堵住了也不是什么大事，因为司机总能够在四通八达的网络里找到其他路径，而经验丰富的老司机甚至能够通过出入两条相同走向的大街中间的停车场抄近路。这样的绕行虽然拖延了时间，但是不至于让人止步不前。

敲除某个与代谢有关的基因就像堵死了某一条主路，阻碍了代谢的原料进入生化反应的错综网络。而一条备用的代谢通路就是一条可以迂回的支路，位于断点位置后方的生化反应很快就会消耗完先前积累的分子，所以生物体需要绕过原先的通路，找到一个合成原料的支路继续反应，确保生命在代谢的城市里畅行无阻。这可不仅仅是一个抽象的比喻而已。生物工程学家能够用敲除代谢相关基因的方式阻断特定的通路，而当他们这么做时，生物体往往能够重新分配原料物质的走向，保证必需物质的合成从而存活下去。对于代谢而言，支路反应的存在甚至比单纯的后备基因更重要。

当然，发育稳态不仅仅局限在代谢或是基因组的水平。在单个蛋白质中这种现象同样普遍存在。比如溶菌酶（lysozyme），这种蛋白质通过摧毁可以保护细菌的细胞壁杀死它们。溶菌酶不仅存在于人类的唾液、眼泪和母乳中，而且还广泛存在于许多其他动物体内，某些攻击细菌的病毒也包含溶菌酶。当科学家想弄清楚这种蛋白质如何工作时，他们采用的方法类似于从基因组中敲除基因的研究，只是动作要小得多：他们改变蛋白质内氨基酸链上的一个氨基酸，然后观察这种改变的后果。

在获得超过2 000种溶菌酶的变种之后，他们发现其中大约有1 600种依旧能够杀死细菌，每一种变种内都只有一个氨基酸发生了改变。也就是超过80%的溶菌酶变种在更改了某个氨基酸之后，依旧具有杀菌作用。以溶菌酶为代表的蛋白质，就像代谢一样，具有很强的发育稳态。同样的情况也适用于调控环路——我们已经提过大肠杆菌的调控环路在实验室中经过大刀阔斧的改造后功能依旧不发生改变的情况了。

发育稳态最明显的优势在于保证生物的生存。这种作用可以追溯到

第一个能够自我复制的RNA出现，微小的复制错误会在RNA传代中不断积累直到复制无法进行，而发育稳态则能够帮助RNA对抗致命的复制错误。这是现实版的“第二十二条军规”[19]：RNA分子必须在复制中尽量减少错误来保持自己在复制中不出错的能力。不过现代的RNA只需要些许的发育稳态就能显著降低复制错误的发生率：因为些许错误很难改变这种稳定的分子的自我复制能力，发育稳态为RNA分子发生复制错误后提供了稳定复制的喘息时间，而在这段时间内可能会有更好的自我复制分子横空出世。

发育稳态的重要性远远不止于此，它还可以用于解释基因型网络和进化的动力。

让我们重新回到我们拜访过的自然图书馆，在那里，每个代谢（每种蛋白质或者每个调控环路）都与某一馆内的馆藏相对应，而与每一本馆藏相邻的馆藏都只与它相差一个字母，这个不同的字母可以是一个生化反应，也可以是一个酶或编码这个酶的基因。我们从敲除基因的实验里得知，例如通过基因敲除阻断某个代谢反应，许多这样的基因修改都不会对生物造成可见的不良后果。也就是说，至少对于生物体直观可见的特征而言，即便基因型发生变化，生物的表现型也不一定随之改变，我们称这样的现象为发育稳态。而发育稳态的强度则由馆藏所在的社区大小来衡量，即每次只促成一个微小的改变而表现型保持不变的可能性大小。

社区里相同表现型的馆藏越多，生物的发育稳态就越强。如果我们假设生物不具有这种发育稳态，你就能看到它带给生物的优势：如果一种代谢、一个蛋白质或者一个调控环路没有任何邻里，它和它控制的表现型将极度孤立无援，弱不禁风。而在另一个极端上，如果所有微小的改变都能够保持表现型的一致，那么发育稳态的强度将达到峰值：任何单独的微小变异都不能改变这种性状。

无论是极端脆弱还是极端稳定，在现实世界中都不存在。没有哪个生物的哪个性状是绝对脆弱的，也没有哪个性状是绝对稳定的。所有的生物，它们所具有的结构和行为，都在一定程度上具有发育稳态。正是这种稳态赋予了不同生物种群探索巨大的自然图书馆的能力。图书馆中相同主题和内容的馆藏数量惊人，但每个主题的馆藏都仅仅是图书馆内无数书籍中的一小部分而已，自然图书馆就像海洋，相同主题的馆藏仅仅是海洋里的一滴水，而每个文本不过是组成这滴水的水分子而已。

如果没有发育稳态，图书馆中依旧可能会有许多主题相同的书，但

是它们相互之间都会相距甚远，毫无瓜葛。没有读者能够在一本书附近的馆藏里，找到探讨类似主题的书，它们之间往往只相差一页、一个单词甚至一个字母。相同表现型的不同基因型就像在夜空中眨着眼睛的星星，相互之间隔着以光年计的寂寥空间。

值得庆幸的是，生命的世界并不像宇宙。以任何一个文本作为起点，我们都可以选择这个文本的众多表意相同的相邻文本，再以同样的方式行进到下一个同样稳定的文本上，如此反复，而不改变我们需要的那个主题。用这种方式，生物体能够探索自然图书馆里未知的领域，并有机会发现新的性状。发育稳态使得表现型相同而基因型不同成为可能。由此，大自然可以在发育稳态所创建的基因型网络里琢磨新的乐高积木拼法。

我们在第2章中提到过自组织系统，自组织的原则在生物界和非生物界都普遍存在，从星系的形成到生物膜的组装都涉及自组织，而基因型网络恰好就是这种系统的又一个例子，但它又是比较特殊的那一个。基因型网络和星系不同，星系的形成依赖宇宙空间中物质之间的引力作用，而生物膜的自动组装则有赖于脂质分子与水分子之间的亲疏关系，但是基因型网络并不会随着时间的流逝而改变，它们是格局无限大的自然图书馆中的常住居民。

即便如此，基因型网络具有某种组织性的事实仍毋庸置疑。由于它的组织形式复杂，我们到今天也不过是略知皮毛而已。但是我们可以肯定，基因型网络具有自组织性。与星系形成和生物膜组装相比，基因型网络自组织性背后的原因要简单得多：因为生命具有发育稳态。对于基因型网络而言，发育稳态一方面不可或缺，否则表现型相同的不同基因型将被孤立；另一方面，发育稳态也别无他求。只要代谢、蛋白质和调控环路具有一定的发育稳态，基因型网络就可以在宇宙中生根发芽。



发育稳态足以维持基因型网络的存在，但是仅有基因型网络对进化而言还远远不够。原因在于，进化的发生必须要同时满足两个看起来相互矛盾的条件。进化需要生物同时具有保守性和可变性。就像当初那些企图横渡大西洋的先锋飞行员一样，他们也需要参考莱特兄弟的飞机原型：他们要的是能够完成这项壮举的新飞行器，但是他们同样需要学习如何让不够完美的旧飞行器在天上翱翔，直到新一代飞行器取而代之。

同样的道理在自然界也适用，大自然需要保证生物的存活，同时寻求新的性状。基因型网络为探索新性状提供了便利，但是网络本身对保留已有的性状并没有什么助益。

我们需要再次强调这一点，因为基因型网络的发现让我们在惊喜兴奋之余，也容易冲昏我们的头脑，忘记自然选择所扮演的重要角色。自然选择的作用体现在它的保守性，它是进化的记忆，保留了所有值得保留的改进，无论改进多么微不足道，假以时日，这些微小的改变终会积流成河，聚木成林。我这样说也是有据可依的。达尔文在他的物种起源中有一段关于眼睛的描写，眼睛无疑是生物进化史上最出色的成果之一。“眼球精密的结构是无与伦比的，它能够调节焦距以适应不同距离，能够调节进出的光量，能够纠正色差和球面像差，我必须坦陈，造就这一切的自然选择对我而言简直不可思议。”

当光线穿过我们的眼球，眼球中的棱镜系统就把外面的世界清晰地投射到了可以感应光线的视网膜上。这听起来简单，不过在这个过程中，眼球必须以精密的角度改变光线前进的路线。对光进行折射可不仅仅是改变晶状体的形状就能做到的，构成晶状体本身的材料常常被忽略，而它对成像而言至关重要。晶状体的组成成分由来已久，它的出现依赖于新的调控通路。

向水的表面投射一束光线，你能够在水面与空气交界的地方看到光线的弯折。如果在水里溶入糖，那么光线弯折的角度会变得更小[20]。溶解的糖越多，折角越小。食品工业正是利用这个原理检测酒、软饮和果汁中的糖含量。我们的眼睛也利用相同的原理对光线进行折射，区别在于眼睛利用的是蛋白质而不是糖。这种蛋白质（晶状体蛋白）在晶状体中有极高的浓度，使得晶状体对光线具有极强的折射能力。

晶状体蛋白在光线折射方面的作用惊人的出色，很容易让人以为它的存在是为晶状体的形成而量身定制的。然而，事实并非如此。许多晶状体蛋白都是参与代谢的酶，除了数量上相对较少以外，它们和身体中参与其他生化反应的酶相比，并没有明显的特殊之处。不同的生物利用不同的酶作为晶状体的蛋白成分。这些酶与其他蛋白质的一个不同在于，它们不容易凝结成块，哪怕在眼球内以高浓度的形式聚集也不会轻易析出。

眼球利用的蛋白质本质上是一种酶，但它不要求这些蛋白质能够像乙醇酶一样分解酒精，只因为它们是透明的，就像你找了一块破砖头做书立不过是因为它碰巧很沉而已。此外，晶状体蛋白往往非常稳定坚

韧，人类眼球中构成晶状体的晶体蛋白往往将伴随一个人的一生，从出生直到去世。但是有的时候晶体蛋白也会出问题并凝结析出，使晶状体呈现乳白色。这就是我们所说的白内障，白内障最终必然致盲的结果人尽皆知。

达尔文本人无缘得知任何关于蛋白质化学领域的知识，但是他做出了大胆的假设，他认为脊椎动物美丽的眼睛，以及眼睛内精妙复杂的晶状体，都是一系列微小进化积累的最终产物，这一洞见已被今天的我们所证实。早在我们的祖先选择不易析出的代谢蛋白作为晶体蛋白之前，它们的祖先（某些蠕虫或是海星）就已经开始利用感光细胞了，这些感光细胞的作用可以帮助它们找到阴影下的藏身之地以躲避掠食者。

数百万年之后，感光细胞逐渐聚集在一个浅浅的凹陷内，形成视杯（eyecups）。视杯能够比感光细胞更好地感受光源的方向，视杯进一步凹陷就成了视坑（pit eyes），视坑对于光源的感应已经相当出色。再进一步，视坑逐渐进化成视孔（pinhole），视孔终于能够真正意义上地进行成像了。到此，眼球的形成只差一种能够折光的高浓度透明组织就完成了，比如晶状体蛋白。此时，距离晶状体的出现只有一步之遥。我们最终在眼睛的结构里看到了能够移动和变形的晶状体，因为有了它，眼球才能呈现清晰的物像。

每一个小小的改变和进步都值得被保留，而自然选择的确也做到了。之所以如此肯定，是因为许多动物身上依旧保留有这些改变：某些扁形动物中还保留着视杯，蜗牛身上仍旧存在视坑，而鹦鹉螺——一种贝壳体分为许多小室的乌贼近亲，身上则有视孔，而水母等动物身上则具有相对简单的原始晶状体。

在气势宏伟的中世纪大教堂中，教堂里高耸的尖顶和巨石雕琢的厚重圆柱都会配以无比精致的装饰，高高的拱形屋顶往往超出了我们的视线范围，掩映在半明半暗之中。而由所有这些细节构成的最终成品，如果没有人告诉你它们是一块砖一块砖修筑起来的，你也许很难相信世间竟能有此等杰作。而我们的眼睛也是这样的杰作。

分子进化的过程亦是如此。北极鳕鱼体内的抗冻蛋白可不是像雅典娜[21]那样在一夜之间就形成的。北极鳕鱼祖先体内的某种蛋白质以一次一个氨基酸的速度缓慢积累着有益的变异，每次变异只要把体内液体的凝固点降低仅仅0.1摄氏度，其后代的生活范围就可以向外扩展数公里。更大的生存空间也意味着数量更多和种类更丰富的食物供给。只要是能带来生存优势的突变就具有保留价值，而一系列类似变异的积累则把鳕

鱼的生存极限延伸到了极度低温的疆域里。基因型网络对于寻找新性状至关重要，而自然选择则是新性状的保留者。



通过数量积累逐步改良生物性状的突变对生物进化来说至关重要，不过这并不是DNA改变的唯一方式。许多突变在首次出现时对生物往往既没有好处也没有坏处。这种中性的突变需要归功于发育稳态，发育稳态使得生物体对错误有了一定的耐受性。

中性突变对进化或许也很重要，至于为什么，我们至今仍不甚明了。事实上，有关自然选择与中性突变的争论由来已久，在20世纪的最后30多年里，中性突变一直是达尔文主义者的眼中钉、肉中刺。分子生物学领域的技术革命在自然界的生物体中发现了数量惊人的基因多样性，从哺乳动物到果蝇，甚至微观的细菌：同一个物种体内的数千个基因往往在不同个体之间存在DNA序列上的差异。多数科学家相信，这些多样的基因都是自然选择的结果，而持有这些看法的科学家往往是忠诚的达尔文主义者。他们认为，既然多样的基因得以保留，说明同一个基因的多样化有助于生物的生存和繁殖。

但是这些自然选择论者却遭到了一个来自少数派的微弱的反对声音，他们就是中性主义者。中性主义者认为，大多数的突变对生物体没有任何改变，因此也不会受到自然选择的影响。至少在这些突变首次出现时，它们是绝对中性的。在某些科学家眼中，比如古生物学家史蒂芬·杰伊·古尔德（Stephen Jay Gould），中性突变的存在极大削弱了自然选择在生物进化和新性状形成中的作用。

科学和技术的发展史为中性突变提供了一个不甚恰当的类比对象，即新理论和技术发明之初往往找不到自己的定位，但是它们的价值有可能在未来某个时刻变得不可估量。数论就是一个很好的例子。数论是数学的一个分支，美国数学家伦纳德·迪克森（Leonard Dickson）曾经不止一次说过："感谢上苍，数论没有被任何应用技术所玷污。"从欧几里得开始直到1919年，这句话都没有什么问题。但是在仅仅几十年之后，数论就因为一个完全不相干的领域而被放在了互联网经济的中心位置。

随着数字计算机和互联网技术的发展，数论被用于网络信息的加密，为电子金融和网上银行业务保驾护航。类似的例子还有很多，德国物理学家海因里希·赫兹（Heinrich Hertz）通过实验验证了由詹姆斯·克

拉克·麦克斯韦提出的电磁理论，但是赫兹本人也说不出这个理论到底有什么应用价值。赫兹曾经不停地跟别人说他的实验“一点儿用都没有”，“只不过证明了大师麦克斯韦的理论是正确的而已”。而在短短40年之后，赫兹的实验促成了世界上第一座获得商业许可的广播站建立，即位于匹兹堡的KDKA，这个无线电台至今仍在1 020千赫的频段上播送节目。

中性学说最著名的支持者莫过于日本科学家木村资生（Motoo Kimura），木村资生建立了一套复杂而成功的数学理论，以评估中性突变在进化上的命运。木村资生主张自然界大部分的遗传突变都是中性的。不过在基因组学时代到来的今天，我们已经知道，木村在这一点上的认识是错误的，中性突变并没有比那些带来优势的突变更多。撇开这点不谈，木村认为中性突变十分重要的观点则是完全正确的，尽管我们又花了数十年时间才理解中性突变之所以重要的原因。

第一个原因是，中性突变在探索基因型网络时非常重要。中性突变为读者在自然图书馆中寻找新性状时铺就了一条安全的道路，避开了一路上诸多无意义的伪作。如果没有基因型网络和中性突变，安全探索自然图书馆几乎是不可能的。

第二个原因是，以中性突变出现的变异不一定永远是中性的。中性突变同样可以在某一天成为有意义的突变，就像曾经毫无应用的数论那样。一旦中性突变表现出有利于生物生存的性状，它们就会被自然选择保留下来。换句话说，自然选择学说和中性学说都没有错，自然选择和中性突变都是生物进化中的必要组成部分。中性突变先为新性状的出现铺路，而自然选择从众多的中性突变里选出具有进化潜力的那些突变。

一个很好的例子是受到广泛研究的RNA酶，它很好地展示了中性突变和基因型网络能够在多大程度上加快新性状进化的速度。这种特殊的RNA酶作用于RNA链上的特定位置，在所识别的位点上对RNA进行切割。这种核酶的形状犹如它的名字，就像一条双鳍鲨，独特的锤头状是它执行功能的关键——这是一个充分不必要条件。在广阔无边的自然图书馆里，同样催化活性的其他RNA分子很可能有着不同的形状、不同的表现型，从而使得核酶剪切RNA的刀刃更加锋利。

如果没有基因型网络的存在，那么RNA图书馆里的一小簇读者，也就是某一簇进化的RNA分子，必须聚集在“43核苷酸区”里，它的脚步只能延伸到相差一个核苷酸的馆藏里。按照这个范围，某种特定的RNA酶只有129个不同的相邻RNA。由于我们能够计算每一种RNA的形状，所

以我们知道这129个RNA一共有46种不同的形状，这也是进化在不借助基因型网络的前提下能够触及的分子形状数目。

如果我们继续下去呢？如果只选择那些突变是中性的相邻RNA，也就是同样是锤头形状的邻居分子，并检验它们的相邻RNA，这样我们就能得到962种新形状。而后我们再深入一步，检查那些相邻RNA里中性突变的相邻RNA，这次我们将得到1 752种新形状。仅仅在基因型网络中深入两步，新的分子形状数量就比核酶直接相邻的RNA多出了将近40倍。而沿着锤头状这个分子特征前进的基因型网络支路远远不止这两步，这张网络中有超过10^{19}种RNA，现有的计算机还没有能力对所有相邻RNA的分子形状进行推演。不过我们几乎可以肯定，在这些RNA中有数十亿种新的分子形状，而且这些形状不同的分子都具有相同的功能，因为在进化过程中，保证生物的生存本身就是基因型网络存在的意义。

这就是基因型网络加速进化的方式。基因型网络就像科幻小说《星际迷航》中用来进行超光速星际旅行的曲速引擎。如果没有基因型网络的存在，那么锤头状核酶的进化速度将放缓至现有速度的1/40。与已经进化了将近40亿年的现代生命相比，如果没有基因型网络的帮助，地球生命相当于才刚刚走完第一个一亿年的路。在生命出现最初的一亿年里，地球上可能已经有了为数不多的几种细菌，但是多细胞生物肯定还没有登上历史的舞台，更不用说鱼类、陆生植物、恐龙或者人类了。基因型网络对生物进化的加速远远超过40倍，确切的数字我们甚至都无从计算。如果没有这张基因型网络，地球上的生物可能到今天都还没有爬出那盆原始浓汤。

在《星际迷航》中除了曲速引擎，还有一种实现超光速飞行的方式：直接改变宇宙空间。创意无限的科幻作家们创造了许多新奇的技术，比如虫洞（wormhole）。虫洞让数千光年的旅程在一瞬间就能完成，而我们发现基因型网络的作用也与之相似。基因型网络缩短了图书馆内两部馆藏之间的距离，无论这座图书馆是代谢、蛋白质还是调控环路图书馆，基因型网络都发挥了这个作用。

我们来想象有一群参观图书馆的读者，它们是某个物种的一个种群，它们聚集在某个文本周围。这个文本叙述了一个特定的调控环路，而该环路控制的基因表达谱会参与身体某些部位的发育，比如鸟类的翅膀。我们再假设在这座图书馆的某个角落里另有一本书，它叙述的调控环路和基因表达谱，能够让翅膀的空气动力学和重量都得到些微优化。这本书藏得越深，读者想要找到它所花的时间就会越多。

在规模如此惊人的图书馆里寻找某本特定的书犹如大海捞针。你当然可以找到那本书，不过前提可能是你要走遍海里的大部分地方，可能你要走遍每一个角落才能成功。只不过相比某个特定的表达谱而言，图书馆的规模更像是宇宙而不是大海。常识告诉我们，要捞起这根针，我们将不得不走遍宇宙的每个角落。

但是自然图书馆从来不落俗套。这一点在我们发现不同的调控环路能够控制相同的基因表达谱时就已经略有涉及了。我们知道，大海里的针往往不止一根。而当我们在图书馆里寻找特定表达谱的调控环路时，却发现自然图书馆比我们原先认为的更加怪异。在这项研究中，我们首先以随机的方式设计出数千种基因表达谱，而后用计算机针对每一个表达谱生成一对调控环路。两个环路中的其中一个与目标表达谱对应，而另一个则不然。除了所控制的表达谱，每对环路内基因之间的作用关系也不尽相同。接下来，我们逐渐改变第一个调控环路里的基因的作用关系，一次改变一个并保证环路所控制的基因表达谱不变。

经过足够多次的操作，第一个环路能够与配对的环路变得类似吗？答案是，可以。我们的研究发现，对于含有20个基因的调控环路来说，配对环路中基因相互作用形式的相似性能够达到85%。换句话说，以图书馆内任何位置作为起点，你不需要长途跋涉，只要脱离基因型网络走出去15步就能到达另一个基因型网络。也就是说，无论你从大海里的哪个位置开始寻找，你要找的针总是在你周围。

如果这还不够让你觉得新奇，那么我们来看看更诡异的东西。

我们假设自然图书馆就是图6-1中的正方形，正方形中的黑点是某个文本。围绕黑点的圆圈，它的半径相当于正方形边长的15%。这是读者要从一种表达谱到达某个新基因型网络的平均距离，15%这个数值来源于我们上述所做的研究。我们来做一个简单的算数，如果圆的半径是15厘米，那么正方形的边长就是100厘米，圆的面积约为707平方厘米，差不多相当于正方形总面积的7%。

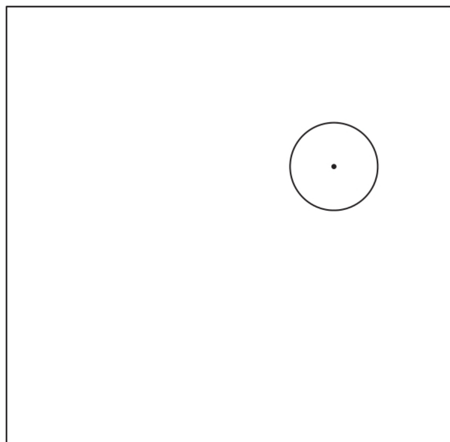


图6-1　自然图书馆与基因文本关系示意图

当然，现实生活中的图书馆不是二维的，它们都存在于三维空间。出于简化考虑，我们把图书馆抽象为一个立方体，那么在这种情况下，一个表达谱所在的社区就相当于一个球。球的半径依旧根据立方体边长的15%设定，但是两者的体积之比却发生了变化。球的体积不再是立方体体积的7%了，而是1.4%。

而调控环路图书馆甚至连三维、四维图书馆都不是。它们位于更高维度的空间，在那里，图书馆是超立方体，社区则是超球体。在四维空间中，超球体是超立方体体积的0.2%。而五维空间中，超球体是超立方体体积的0.04%。

在调控环路所处的高维空间中，这个比例超出你的心理预期。超球体与超立方体的体积比不是0.1%，也不是0.01%。而是仅有10^{-100}%。对于图书馆里的读者而言，要从一个文本出发寻找新的基因型网络，只需要探索一块微不足道的区域。高维空间体积比例不断缩小源于一个简单的几何规律：越高维的空间内，半径为边长恒定比例的超球体在超立方体中所占的比例总是越小。体积比例的下降不是因为我们所举的例子中的半径边长比为15%，不管这个比例为多少，哪怕是高达75%，三维空间中球体与立方体的比例也会降为49%，在四维空间中降为28%，在五维空间中进一步降为14.7%。随着维度升高，比例越来越小。

其他图书馆里同样存在这种反直觉的现象：图书馆所处的维度越高，也就是说，馆内的代谢和分子种类越多，找到新性状的难度也就越小。生物一旦在自然界站稳脚跟，想要再寻找新的性状并不需要花费太多力气，它们只要改变少数几个生化反应，探索代谢图书馆的一小块区

域，就可以撞见它们所期盼的新性状。对于RNA而言也是同样的道理。以一个已有的RNA作为起点，你只要稍稍改变它的核苷酸就能够获得形状不同的新分子。在图书馆里寻找新性状的代价不过是回身走两步而已。

适者降临的代价仅仅是探索自然图书馆的$10^{-100}\%$，如此看来，自然界依靠略显盲目的探索方式却依然能够造就生物丰富的多样性也就不奇怪了。进化不用遍寻海底的每个角落，因为遗落在海底的针远远不止一根。事实上，海底散落的针都在一张紧密编织的网里，而这一切都要归功于发育稳态和它对基因变异的耐受性。

不知道你是不是已经有了这样的印象：图书馆中每个文本所在的社区规模都极其庞大。那么你大概也很容易理解，图书馆在构建组织上的另一个特征：每个基因型网络不但分布广泛，而且不同的基因型网络之间还有频繁的交织互动。它们紧密交联，互相围绕，繁复多变。整个网络由数百万乃至数十亿根丝线织成，每根都对应一种特定的表现型。如果给每根丝线涂上一种独特的颜色，那么这张复杂无比的网络中的每一条丝线上都有数十亿根其他颜色的丝线交织而过。如此精致的绸缎只可能存在于高维空间中，它的华美与我们生活中的任何织物都不同，完全超出我们的想象。这张网隐藏在我们生活中的每个生命体背后，生命由它而生。



复杂交联的基因型网络都是发育稳态的产物，发育稳态对于进化而言至关重要。不过，天下没有免费的午餐，发育稳态也一样。发育稳态的代价就是它的复杂性。

想要声讨某个事物太过“复杂”似乎没有必要。在路易斯·卡罗尔的《爱丽丝梦游仙境2：镜中奇遇记》中，爱丽丝在探索一块神奇的棋盘时遭到了红衣骑士的攻击。所幸爱丽丝被红衣骑士的死对头白衣骑士及时救下。但没想到的是，白衣骑士是一名发明家，他迫不及待地要向爱丽丝展示他的发明，比如一个只要按一下底下的按钮就可以遮风挡雨的箱子，一个马背上用的捕鼠器，还有一个用吸墨纸、封蜡和火药做的甜点。

“你瞧啊，”白衣骑士稍稍停顿了一下，然后继续说，“所谓

有备无患嘛。看看我的马戴的那些脚镯子。”

“它们有什么用？”爱丽丝饶有兴趣地问。

“以防鲨鱼的袭击，”骑士回答，“这是我的专利。”

白衣骑士深受复杂而奇葩的发明掣肘，沉重的累赘让他无法策马奔腾，无法陪伴爱丽丝继续她的旅程，理所当然地，白衣骑士很快就在故事里谢幕了。不过，他作为“大道至简”的反面典型，还一直活在读者的心中。

早在卡罗尔写完爱丽丝的故事之前，14世纪英国修道士“奥卡姆的威廉”（William of Ockham）一直都是精简主义的狂热信徒，他创立的精简原则在今天已经被视为简约性的定义：万物本质当由最少的事实作为支撑，切勿浪费。奥卡姆也把事实称为“实体”（entities），这个观点也被世人叫作“奥卡姆剃刀”（Ockham's Razor，也作Occam's Razor），它让人相信普世的科学理论往往形式简约。此外，奥卡姆剃刀也是工程师制造机械时的追求，虽然他们在工程学中已经有了更接地气的座右铭：KISS，即“还可以更精简，你个蠢货”（Keep it simple, stupid）。

奥卡姆剃刀不仅是出于美学或者哲学层面的追求，在工程学里，它还有着经济方面的考虑。量产一台机器的每个零件都需要成本，更少的零件意味着更低的成本，而降低成本是每个工厂老板都希望看到的。另外，装配过程复杂的机器也更容易出现安装错误。对于制造一台机器而言，精简主义大有裨益。

所有试图理解生命，却对它们的复杂性望而却步的生物学家，大概都会对精简主义心向往之。生命在很多方面似乎都复杂得没有必要。调节昆虫体节分化为14段的调控环路中有数十种分子，不过科学家从很多年前就了解到，只需要这数十种分子中的两种就可以实现同样的功能。昆虫体节分化的原理研究耗费了人类数十年的时间，自视甚高的人类工程师也只能对这些小虫子甘拜下风。不知道你是否还记得由生化代谢构成的交通网络，里面布满了备用车道、迂回路线以及平时不太常用的小巷子，以上这些现象都有一个相同的问题：为什么它们会存在？为什么优胜劣汰、效率为先的大自然会保留这些多余的复杂性呢？

答案是“环境”，确切地说，是“各色环境”。看似浪费的复杂性，实际上却是基因为了应对各种不同的环境留下的后手。

在营养极度贫瘠的环境里，大肠杆菌只能利用单一的碳源物质合成

自身所需的必需成分，比如氨基酸和脱氧核苷酸。在这种情况下，大肠杆菌体内3/4的代谢反应都是无用的累赘。就算把它们都敲除，大肠杆菌照样能够生存，这就是发育稳态。

然而环境是复杂多变的。如果单一的碳源从葡萄糖变成了乙醇，原本“无用”的生化反应就成了大肠杆菌生命延续的关键。大肠杆菌能够利用80多种碳源合成必需成分，这些生化反应大多布局精巧。碳元素仅仅是生物所需的众多必需元素之一，其他元素的代谢需要进一步的生化过程。代谢反应类型的多样化有利于生物在多样的环境中生存。对于生物而言，高度的复杂性也就意味着对不同环境的高度适应性。

基因组中保留的重复基因同样是为了应对多变的环境。重复基因在诞生之初是完全相同的，不过它们不久之后就会踏上不同的命运之旅。突变会在各个基因内不断积累，改变它们的DNA序列和基因表现型，以便能够应对特定的环境。在人体内，某些催化分解反应的酶在肝脏中活性最强，而与之同源的另一些酶的最适宜化学环境则是在大脑中。而真菌中的一种在葡萄糖充裕时负责把糖分子转运入细胞的蛋白质，它的另一种同源蛋白则在葡萄糖稀缺时负责清除糖分子。许多重复基因的实际作用依旧是个谜，也许它们是在为某些还未遇到的特殊状况而蛰伏。

我们在工程技术领域也能够找到类似的例子。虽然工程师们对于精简主义相当推崇，但他们同时也要为多变的环境做足准备。比如，如果你只是想顺着河水随波逐流，那么你只要有一艘木筏足矣。如果你想去河对岸，那么你需要更复杂的设计以便控制木筏的前进方向，这时你又需要一个船舵。如果你不想被浪头打湿，你就需要在木筏上加装船身。而如果你想要逆流而上，就得给船装上船桨或船帆，哪怕是腓尼基人和埃及人在5 000年前就发明的横帆。原始归原始，横帆在船只顺风的情况下确实非常实用。相对而言，当风向改变时横帆就略显欠缺，而逆风情况下则会彻底沦为鸡肋。此时你就需要双纵帆：一面位于桅杆前方的三角帆和一面紧随其后的主帆。为了应对变幻莫测的暗流、大浪和风向，略显复杂的工程设备必不可少。

反过来道理也一样。至少在生物学领域，如果随着时间的推移环境一成不变，那么发育稳态相对而言就不那么重要了，遗传的复杂性也就会随之降低。对这一点我们不需要挖空心思，只需要看看家里养的盆栽就能够找到鲜活的例子。确切地说，这个例子是生长在植物上的某种昆虫。

蚜虫（Aphids），也叫木虱、腻虫或蜜虫[22]，数千年来一直是农民

和园丁不共戴天的死对头，尽管在总计4 000多种已知的蚜虫中只有100多种会吸食农作物的汁液。除了家里的景观植物，蚜虫的食物也包括棉花、各种果树以及谷类作物。蚜虫和19世纪40年代的爱尔兰大饥荒[23]以及50年代的法国葡萄酒庄园大虫害[24]也脱不了干系。蚜虫是最具破坏力的农作物寄生虫之一，但就其在科学研究中的地位而言，蚜虫的价值不可估量。蚜虫体内隐藏着一种更小的生物，它为我们上了有关发育稳态和遗传复杂性的宝贵一课。

许多人都知道蚜虫依靠吸食植物的汁液生存，但很少有人知道的是，植物汁液的营养并不丰富。植物汁液缺乏某些必需的物质成分，包括几种生物必需的氨基酸。为了弥补这个缺陷，蚜虫和一种大肠埃希氏菌的近亲组成了搭档。这种学名为Buchneraaphidicola的细菌主要栖息在蚜虫体内。

蚜虫和体内细菌的同盟关系使得它们可以同时受益，这种关系又被称为内互利共生，是一种高度亲密的共生关系。蚜虫体内的细菌不光栖息在蚜虫体内，而且直接栖息在蚜虫的细胞内。为了支付“房租”，它们为宿主细胞提供了救命的物质：合成必需的营养分子，尤其是蚜虫本身不能合成且在植物汁液中也无法摄取的必需氨基酸。对于蚜虫而言，体内的细菌就如同延续自身生命的工厂。

鉴于细菌的汗马功劳，蚜虫也会投桃报李。栖息在蚜虫细胞内的细菌简直就像漂游在一碗肉汤里，任何食物都伸手取用即可。除了食物，蚜虫细胞还为细菌提供了安全舒适的庇护所。身携细菌到处游走的蚜虫可以为细菌遮风挡雨，御寒保暖。与蚜虫共生的细菌不需要担心作物歉收，不用提防掠食者或者其他威胁，它们只要兢兢业业为宿主服务就能衣食无忧。蚜虫体内的细菌犹如一个与世隔绝的度假者，悠闲地徜徉在大海里，享受着阳光和沙滩，任凭一阵阵温柔的波浪晃动自己的身躯，消磨无聊的时光。

蚜虫体内共生菌的悠闲假期已经持续很久了。宿主和共生菌的首次碰面发生在大约一亿年前，两者一拍即合，从此以后便形影不离。经过一亿年的相伴，有人可能会猜测两种生物都发生了翻天覆地的变化。内生细菌的确变化惊人，通过研究它的进化我们还窥探到了些许发育稳态和遗传复杂性之间的关系。

我们将蚜虫内共生菌与它的近亲大肠埃希氏菌进行了一番比较。大肠埃希氏菌是代谢可塑性领域的大师，它能够利用各种不同的物质作为食物，对多变的环境有着极强的适应能力。大肠埃希氏菌复杂的代谢网

络中包含了1 000多种化学反应，它们使得细菌在应对难以预料的环境时得心应手。

蚜虫体内共生菌的祖先曾经和大肠埃希氏菌不分伯仲。但是好景不长，如今它们的代谢网络中只剩下263种代谢反应。共生菌与蚜虫的同盟关系从恐龙还在陆地上行走的年代就已经开始了，从那时起，共生菌逐渐丧失了自身3/4的代谢反应，而这些反应中的大多数在大肠埃希氏菌中则得以保留。DNA复制错误的持续积累逐渐侵蚀代谢的基因，多数基因都没能挺过DNA的自我移除。而共生菌在无数次基因自我移除后仍旧幸存了下来。

蚜虫内共生菌能够幸存的原因几乎显而易见。多数的基因和代谢反应对于共生菌而言显得无用而多余，与大肠埃希氏菌不同，蚜虫内共生菌所生活的环境在过去一亿年中几乎没有发生任何改变。无论蚜虫如何用尽浑身解数，在不断改变的环境里挣扎求生，共生菌始终沐浴在单调的营养肉汤里。在如此单一的环境里生存，只要一种代谢模式就足够了。在这里，遗传上的复杂性不仅显得多余，甚至可以说是浪费。

蚜虫内共生菌非常特殊，但绝对不是绝无仅有，还有许多微生物也栖息在体型更大的生物体表或体内。它们中有的与宿主互帮互助，而有的则对宿主巧取豪夺。一个著名的例子就是人类体内的肺炎支原体（mycoplasma pneumoniae），它是轻度肺炎[25]的病原体，轻度肺炎患者通常不需要卧床休息。支原体在人体之间传播，依靠人类细胞获得食物，而它的代谢系统甚至比蚜虫体内的共生菌更简单：在营养充沛的人类细胞内，只拥有189种生化反应的支原体依然能够繁衍生息。

让人难以置信的是，支原体的代谢系统中包含了代谢世界中最古老的核心反应：三羧酸循环。此外，极端的精简性也让衣原体因祸得福，它不会畏惧以细胞壁合成酶为目标的抗生素，因为它早就连细胞壁都不合成了。衣原体甚至还会盗用人类细胞的生物膜，包裹并保护自身的细胞内容物。

与精简性相伴的则是发育稳态的降低：不仅仅是对突变，还有对于多变环境的稳定性，两者并非相互独立。对于基因敲除耐受的代谢方式，在多变的环境中同样较为稳定。如果大肠杆菌被限定生活在单一的环境里，例如只有葡萄糖作为唯一碳源，它只需要代谢系统中30%的生化反应就能够生存。但是蚜虫内共生菌则不同，它必须保证263种生化反应中超过90%的生化反应运作正常才能维持生存。只要抹除其中的某一个生化反应就足以抹杀这种生物。

从另一个角度来看，大肠杆菌的代谢网络里有很多备用的通路，而蚜虫内共生菌则没有，它的代谢系统就像一条没有岔路的单行道。只要在路上的某一处设置一个路障，整条路就会被堵得水泄不通。对于必需分子的合成而言，大肠杆菌对于DNA变异和环境改变都相当耐受，而蚜虫内共生菌则不然。

大肠杆菌和蚜虫内共生菌充其量只是代谢图书馆里的两粒飞尘而已，在它们身上适用的规律，即越是适应多变环境的生物在构造上越是复杂，在遗传上越稳定，可能并不是适用于所有生物的普遍法则。我们没有办法在实验室检验所有物种的代谢，不过依旧可以通过计算机演算相当数量的物种，这个研究的原理类似于民调：通过一个较小的随机样本反映一个较大整体的性质。我们选取一个随机样本，将不同的物种代谢置于不稳定的环境中并观察结果，就能够知道大肠杆菌和蚜虫内共生菌到底是自然界的代表还是奇葩。

为此，我所在实验室内的研究人员以尽可能精简的生化反应构建了数百个代谢网络，同时保证不影响生物的生存。我们把这种代谢网络称为最低代谢（minimal metabolisms），对最低代谢进行任何压缩都会导致生命无以为继。我们构建出的最低代谢有的能够在单一环境中生存，有的能在两种、三种甚至数十种环境中生存，每种化学环境之间的差距仅仅在于营养物质的种类。

这项研究的一个直接结论是，通常情况下，在不同环境中生存需要生物具有一定的复杂性。在某个实验中，我们研究了两种以不同物质作为硫源的环境，硫是一种和碳同样重要的元素。我们首先构建了最低代谢，能够保证生物在单一硫源环境中生存的最低代谢远远不止一种，这种代谢体系内只需要不到20个生化反应。而如果要支持生物在5种不同的硫源中生存，代谢体系内则至少需要25个生化反应。当硫源数量达到40种时，最低代谢的容量则扩展到了60个生化反应。换句话说，能够应付的环境类型越多，代谢体系内包含的生化反应就越多，代谢就越复杂。

这种情况下，代谢的发育稳态也会变得越强：我们能够从代谢体系中移除而不影响生物生存的反应相应也增多。代谢体系中的反应数量越多，在某个特定环境中不会用到的反应也就越多。“无用”的生化反应在某种环境中是中性的，但是在另一个环境中说不定就是不可或缺的。大肠杆菌和蚜虫内共生菌并不是特例，它们只不过是一个普遍规律的两个典型：生物的复杂性和遗传稳定性随着它所面对的环境多变程度的上升

而上升。

至此，我们的认识越来越丰富。多变的环境催生了生物的复杂性，而复杂性促成了发育稳态，发育稳态继而造就了基因型网络，而基因型网络的存在让进化成为可能，使得生物能够通过演变适应环境的变化、提高自身的复杂性，循环往复，生物进化通过这种方式螺旋上升。这种进化方式的核心在于处在多维空间的基因型网络的自组织性。自组织性是生命绚烂光彩背后的支持者，它是隐藏的生命建筑师。

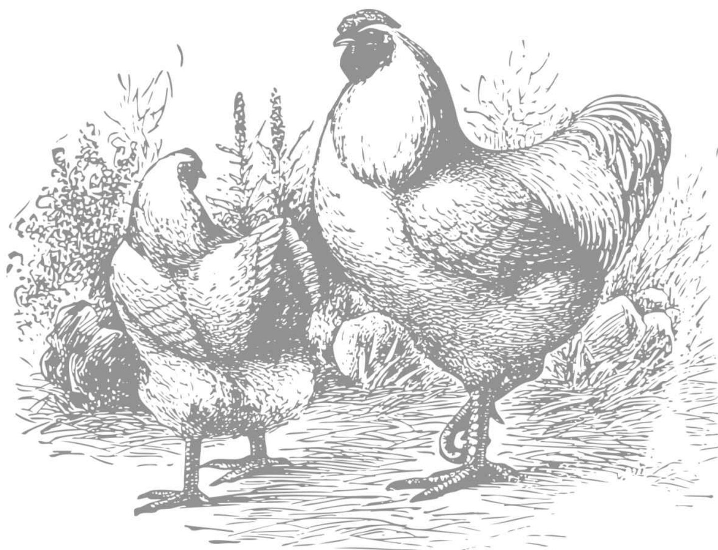


ARRIVAL OF THE FITTEST

07 从大自然到工程技术

Solving Evolution's Greatest Puzzle

自然进化和技术创新有诸多共同之处，促进自然进化的基因型网络在人类技术进步中同样存在。与自然界类似，科研人员也总是行进在各自领域的最前线，他们依赖不断的试错、人海战术、多起源策略和组合优化，模仿自然的创造能力，实现技术突破和创新。技术发明的精简主义和高雅主义，深深隐藏在现实世界的背后。



YaMoR是一种仍处于实验阶段的新式模块机器人，组块之间以铰链相互连接。它既能组成类似千足虫的直线结构，蠕动前行；也能模仿两栖动物的四肢结构，匍匐前进；甚至可以模拟昆虫爬行。模块中的电脑芯片让它能够被重复编程，每一个模块都是智能的，类似于人类的大脑。只要方法得当，模块机器人就可以通过优化自身的设计应对不同的问题。

作为世界顶尖的理工学院之一，瑞士洛桑联邦理工学院设计生产的YaMoR是蒸蒸日上的模块机器人家族中的新成员。目前世界上许多工程学实验室都专注于研制模块机器人，种类众多。但由于适用的条件不同，这些机器人往往风格迥异，相互之间缺乏相似性。有的模块机器人像方形的骰子，有的则像金字塔。还有一些由成群的球体构成，另一些则像成串的轮胎。类似YaMoR的模块机器人在外形设计上几乎百无禁忌，各种立体几何造型应有尽有。

在YaMoR机器人跨出第一步的5.4亿年之前，自然界就已经发生过一次形态设计的百家争鸣了，我们把那场生物性状的爆炸式起源称为寒武纪大爆发（cambrian　explosion）。我们今天看到的所有形态的生物都诞生于那场爆发，而有更多的物种则已经消失在历史长河中，比如所有身体分节、没有四肢的古虫动物（vetulicolia）。与古虫动物门生物相比，以YaMoR为代表的模块机器人显得非常原始。如果推动自然进化的曲速引擎同样能够在人类工程技术中大显神威，那么第二次寒武纪大爆发似乎也就值得期待了。

我们将会看到，促进自然进化的基因型网络在人类技术进步中同样存在，并不是什么天马行空的想法。自然进化和技术创新拥有诸多共同点。

大自然和人类技术发展都依赖于不断试错。作为天才发明家的代表人物，爱迪生曾经“尝试不下6　000种不同的植物，才找到世界上最适合制作灯丝的材料”。他最终在无意间发现竹子是解决白炽灯泡灯丝易断问题的最佳答案，对于这段经历，爱迪生自己也感慨万千，其中一条总结是：“我并没有失败那么多次。我不过是成功找到了一万种不适合的材料而已。”这句名言无疑说明了“试错”对于技术发明而言的重要性，而对于生物学也一样。

试错不管是在当下还是在爱迪生生活的年代都意义非凡。著名的编

程语言FORTRAN曾经帮助科学家模拟从原子到星系的运行规律，为人类理解宇宙立下汗马功劳。FORTRAN的创始人之一约翰·巴克斯（John Backus）曾说：“你必须有随时失败的觉悟。你费尽心力，提出许多不同的想法，然后努力求证，但是结果往往会发现它们都行不通。在偶然发现一个可行的办法之前，你要不断重复这个过程。”

诚然，就失败所蒙受的代价而言，生物进化比一个研究灯泡材料的发明家或者研究理论的科学家要惨重得多。以大雁为例，大雁体内的血红蛋白变体不啻于大自然所做的一场生存实验。如果变异的结果是血红蛋白从稀薄的空气中摄取氧气的能力提升，那么成功。但如果结果是血红蛋白失去携氧能力，那么这种变异连同这个个体就将落入万劫不复之地。

理论科学和工程技术中的失败通常不会有致命的危险，但是错误的观点也往往没有那么容易被澄清。著名的天文和天体物理学家弗雷德·霍伊尔爵士（Sir Fred Hoyle）直到2001年去世都拒绝接受宇宙大爆炸理论，不仅如此，他还为流感理论辩护，认为流行性感冒是太阳风减弱时外星病毒侵入地球大气所致。19世纪的开尔文勋爵曾经用热力学定律以及他的基督教信仰作为测算地球年龄的依据，结果得出的结论只有地球实际年龄的数百分之一。

科学和技术的修罗场上从来不缺充斥着错误信念的聪明脑瓜，而且那些人往往至死不悟。量子物理的创始人之一马克斯·普朗克（Max Planck）曾经颇有洞见地表示：“科学理论的胜利往往建立在异见者的坟墓之上，而不是他们的皈依，下一代人成长过程中耳熟能详的理论即他们认为的真理。”科学就像大自然，总是随着丧钟的节拍翩翩起舞。

巧合的是，大自然抵抗致命错误的秘诀之一恰好被技术发明领域所借鉴：人海战术。探索自然图书馆的生物不止一个，同样的道理，每一项重要的发明也不是某个天才孤军奋战的成果。尽管从在浴缸里泡澡的阿基米德到在专利局上班的爱因斯坦，每个伟大的科学家脑海中的世界对大多数人而言都难以想象，但是如同有成群的生物涌入生物进化的各色化学图书馆一样，科学和技术创新的另一个真相是它们的进步需要密集的资源投入。开发FORTRAN需要众人通力合作，而爱迪生在研制灯泡、电话和电报机的时候也借助了数十名助手的协助。

19世纪的工业革命更是某个新兴阶层崛起的结果，这个阶层的成员主要是受过优质教育的技工，而不是衣冠楚楚的贵族科学家。这些手艺人为了提高生产效率发明了一系列新技术，包括蒸汽机和自动织布机。

时至今日，任何新技术的发明，从新款手机到新式药物，抑或新型能源，都需要大量的科学家和工程师经过激烈的竞争和无数次的失败才有可能成功。从这些例子来看，如果没有试错的过程，我们很难想象有更合理的成功方式：参与的人越多，尝试的可能性也就越多，相应成功的概率也就越高。

与自然界类似，技术和科研人员们总是同时行军在各自领域的最前线。提出过“行为榜样”（role model）和“自我应验预言”（self-fulfilling prophecy）概念的美国社会学家罗伯特·默顿（Robert Merton），因他对于科学理论多起源的归纳而被科学界铭记。默顿把同一个理论的不同起源直接称作“多起源”（multiples）。类似的例子举不胜举：气体温度和压力之间的关系在英语国家被称为波义耳定律，而在法语国家则被称为马略特定律，因为罗伯特·波义耳（Robert Boyle）和埃德姆·马略特（Edme Mariotte）分别通过自己的工作独立发现了这个规律。

罗伯特·富尔顿（Robert Fulton）、马奎斯·德朱弗若依（Marquis de Jouffory）和詹姆斯·拉姆齐（James Rumsey）都是蒸汽船的“设计者”；温度计更是至少有6名不同的发明者。更广为人知的例子是微积分，它几乎由牛顿和莱布尼茨两人同时提出；而伊莱沙·格雷（Elisha Gray）则与格拉汉姆·贝尔（Graham Bell）在同一天申请了电话机的专利（尽管最后还是贝尔赢得了法律仲裁的专利权）。

科研以及技术多起源的现象之所以存在，是因为科学问题与自然界的生命一样，同一个问题通常有多种解决途径。最典型的例子也是科学家研究较为充分的一个问题：生物学家称之为碳固定，而工程学家称之为碳脱除，研究的关注点都是如何除去大气中的二氧化碳。生物除碳的中流砥柱是一种植物酶，它能够把二氧化碳连接到一种名为1，5-二磷酸核酮糖的糖分子上。一旦结合，二氧化碳就会成为新分子的一部分并参与接下来的代谢反应，并最终成为植物实体的一部分。这个过程不仅可以促进植物生长，让今天的我们能够用上化石燃料，它也是碳元素在生物圈和非生物圈循环的方式。

不过自然界并不是只有植物拥有固定碳的能力，某些细菌能够用乙酰辅酶A，而另一些则可以利用三羧酸循环中的某些中间代谢产物固定二氧化碳。试图阻止气候变暖的环境工程师也在研发更先进的碳脱除技术，比如利用乙醇胺或氢氧化钠分子固定二氧化碳。

默顿的多起源理论在许多其他例子中也适用。汽车的发动机可以是往复活塞式，也可以是偏心转子式。汽车的燃料燃烧可以靠汽油发动机

的火花塞来触发，也可以用柴油发动机的压缩热来触发。生物可以用灵活的单眼感知光线，也可以借助复杂精致的眼球感知光线。生活在极圈的鱼类体内的抗冻蛋白前身是不同酶的晶体蛋白及高度多样化的携氧球蛋白，都是生物对于同一个问题的多种解决方式。

科学技术和生物创新的另一个共同点是擅长废物再利用，这在技术发明的历史上体现得淋漓尽致。根据作家斯蒂芬·约翰逊（Stephen Johnson）的记述，约翰内斯·古登堡（Johannes Gutenberg）发明印刷术的契机是他从别人那里借来的一台机器，那台机器配有螺旋传送装置，在设计之初是用来给客人上酒的，而古登堡对它做了改装使它成为驱动新媒体运转的高效引擎。用于加热食物的微波炉技术最早来自雷达：一名雷达工程师在微波融化了他口袋里的一条巧克力棒后发现了微波的加热功能，最初投入商业使用的产品名称正是“雷达炉”。

轻质合成材料凯夫拉（Kevlar），发明的初衷是用以替代赛车轮胎上的钢质材料，而现在则大量应用于防弹背心和钢盔的制造。还有一些甚至都称不上“创造”的普通装置也是“旧瓶装新酒”的产物，比如两个锯木架上摆上一块门板就是一张粗糙的桌子；一只靴子也可以当简易的制门器使用；牛奶箱同时也是上好的档案柜，等等。爱迪生说得好：“想要创造，你需要先有想象力和一堆无用的垃圾。”

1982年，古生物学家史蒂芬·杰伊·古尔德和伊丽莎白·弗尔巴（Elizabeth Vrba）把生物学中的这种现象正式命名为“扩展适应”（exaptation）。事实上，达尔文才是这方面的引领者，他在《物种起源》中就提醒读者，“某个针对特定目的构建的器官可能出于适应更多功能的目的而发生改变”。扩展适应最经典的例子是鸟类的羽毛，构成鸟类羽毛的主要成分与构成爬行动物的鳞片的成分相同，是一种被称为角蛋白（keratins）的致密纤维蛋白。羽毛最初的功能很可能是保温和防水，后来才演变（或者说“扩展”）为用于飞行。

扩展适应的现象在分子水平也很常见，例如那些调节羽毛合成的调控分子。其中一种蛋白质叫“音速刺猬”（sonic hedgehog）。没错，正是那款著名的电子游戏的名字。这种蛋白质在人类体内负责控制手指和脊髓的生长，而在鸟类体内则与羽毛的生长有关。一种控制后肢形成的调节蛋白在蝴蝶体内负责控制眼状斑点的发育，某些代谢酶更是直接参与晶状体的形成。

扩展适应的例子是我们要探讨的自然和技术之间的最后一个共同点：创造在一定程度上就是对原有事物的组合优化。我们第一次了解到

这个事实是在探讨代谢时，生物通过对已有化学反应的组合把有毒的五氯苯酚代谢成可利用的能源物质，或者把体内有毒的成分代谢成低毒的尿素。对于蛋白质而言，新的创造来自对已有氨基酸的重新组合。而在调控环路中，新的环路起源于调节因子之间新的相互作用。技术发明上的创造也逃不出优化组合的套路，以航空工业的喷气式发动机为例，它的三个主要组成部分分别是提高气压的空气压缩机、混合空气和燃料的燃烧室以及产生动力的旋转涡轮。

在20世纪，当来自英国、德国、匈牙利和意大利的工程师们埋头研制世界上第一台喷气式发动机时，上述三个部分都已经不是什么新鲜玩意儿了。最早的空气压缩机是装在熔炉上的风扇，人类利用工业风扇的历史已经超过了2 000年；蒸汽机车和内燃机车则是燃烧室不可或缺的操作设备；阿基米德在公元前3世纪就已经发明了螺旋涡轮，第一个燃气涡轮发动机的专利出现在18世纪晚期的英国。

喷气式发动机的发明并不是优化式创新的特例。几十年前，以经济学家约瑟夫·熊彼特（Joseph Schumpeter）和社会学家科拉姆·吉尔菲兰（S.Colum Gilfillan）为代表的社会科学家都认为对已有事物的组合优化是发明创造的关键。经济学家布莱恩·阿瑟（W. Brian Arthur）在他的《技术的本质》[26]中甚至直言，“无论什么新技术都必须建立在已有的技术基础之上”。从前面的章节我们大概能够体会，同样的道理在生物学中也适用：任何生物进化中的新性状，无论它在无尽的宇宙图书馆中的哪个角落，都是组合优化的结果，就像每一本新书都不过是对已有文字的重新组合而已。

试错、人海战术、多起源及组合优化都是科学技术和自然界之间相似的地方，难怪技术学家一直想要模仿自然的创造能力。这里我指的不仅仅是生物技术专家，尽管人类在生物技术领域已然硕果累累：从可以把沾满泥渍的裤子洗得干干净净的含酶洗衣粉到糖尿病患者使用的人工合成胰岛素、通过基因工程培育的抗虫作物。生物技术的材料取自生物本身，因此它从一开始就已经利用了自然图书馆带来的便利。我在这里想搞明白的是，与人造材料打交道的技术学家是否也能享受同样的便利，比如利用玻璃、塑料、硅质材料或YaMoR的专家。

技术创造与生物进化并没有我们想的那么神秘，反而特别因循守旧。技术学家们早就发现，创造的过程犹如一个按部就班的算法，连机器都能重复。变异通过改变DNA创造具有新表现型的生物，其中一些经过自然选择幸存下来并繁衍生息，这个过程就是在变异、选择中不断循

环往复。意识到这一点的技术学家，确切地说是计算机科学家们，据此创立了一个全新的领域，以研究生物进化的算法，他们想要完全依靠计算机解决现实世界中的复杂问题。

一个著名的例子是广为人知的旅行推销员问题（traveling salesman problem），这个数学谜题由爱尔兰数学家威廉·罗恩·哈密尔顿（William Rowan Hamilton）在19世纪中期提出。旅行推销员问题的预设条件并不复杂：一个推销员要出门拜访几个身处遥远城市的潜在客户，每个客户都住在不同的城市。推销员需要乘车或乘坐飞机前往，这意味着旅程中会耗费大量时间。而推销员是个恋家的人，他希望在保证造访客户数量的基础上，每趟出差的时间越短越好。旅行推销员问题讨论的是，找出一条经过所有城市的路线，让推销员能够以最快、最高效的方式完成出差。

这个数学问题的难度远远超出它的表象。如果城市的数量不多，几乎任何人都能轻松设计出最短的旅行线路。但是当需要考虑的城市数量超过十几个时，最优路线的制订就变得异常困难。旅行推销员问题属于计算机科学家口中典型的"非确定、多项式"的计算问题。这类命题是目前最困难的计算问题之一，困难的原因在于问题的解决方案将随着城市数量的增加呈指数级上升。

围绕旅行推销员问题已经有上千篇相关论文发表。关注这些论文的读者并不是推销员，而是计算机芯片的设计师。计算机芯片内包含了数以亿计的元件，电子元件之间通过连接进行数据交换。由于缩短元件之间的连接距离能够在节省电能的同时提升计算速度，所以制订电路元件（"城市"）之间的最短路线自然也就成了芯片设计师们的诉求。为百货公司运送货物的卡车司机、联邦快递以及到社区接送学生的校车对这个问题都不会感到陌生。甚至大黄蜂也需要解决这个问题：一只工蜂每次回巢前可能要"拜访"数百朵花，对它们而言，走太多的冤枉路就意味着负担不起的浪费。

如果考量的"城市"数量在数千座上下，那么以复杂的数学手段为推销员设计最佳路线依然是可行的。虽然这些数学理论复杂高深，但是从名字上却一点看不出来，比如"切割平面法"和"分支定界法"。当城市数量上升到百万级别时，这些方法依旧能够制订出接近完美的路线。不过严谨的数学算法并不是必需的，生物学家们愚钝而盲目的算法同样能够解决问题：首先让计算机随机生成一个路线方案——任何路线都可以，无论它多么低效。然后，由计算机程序对生成的路线进行修改，每次只

改变其中几座城市（在不同的情景中也可以是停留的商铺、学校或花）之间的线路，继而查看新的路线是否比原来的更短。如果路线的确变短了，就选择继续改变后面的新路线。下一步再重复上述过程，再比较。而如果线路没有缩短就放弃新路线，回到原有的方案上。经过足够多的尝试，这种简单的算法同样能够让路线变得越来越短，最终找到的路线就算不是最优解，也是相对最理想的路线之一。

这种进化的算法还被应用在了一些你想不到的地方。比如军事作战计划制订员用这种方式设计无人机在敌方领空的最佳巡航路线，密码编译人员用这种方式为敏感信息加密，基金经理用这种算法预测金融市场的动向。汽车工程师也可以通过优化发动机内燃料注入的时间和压力，调整它的运作，而这种算法不负众望，的确能够提升发动机的燃料效率。需要注意的是，仅仅提高燃料效率并不足以推动发动机设计上的改革。

模拟生物进化的进化算法确实是一种强大的工具，但是似乎还缺点儿什么。它们欠缺生物进化的核心部分：组合优化。大自然是组合优化的一把好手，而原因非常简单：标准化。

我们在第2章中已经探讨过，类似通用能量载体三磷酸腺苷以及通用遗传密码等标准化物质的存在，是生命有着共同起源的最佳证据。而人类技术工程中缺少如此高度的标准化，往往只能通过不断创造新的标准推动技术进步，这又构成了组合优化难以实现的主要瓶颈。为了让空气压缩机、燃烧室和涡轮引擎组合到一起推动数吨钢铁飞上天，技术人员着实为不同的工业标准耗费了一番心力。现今汽车内燃机的发明也是一样，发动机的零件接口必须做到标准化，比如活塞、阀门与汽缸的尺寸必须相符。工业革命中许多标志性的发明也经历了类似的过程。第一台投入使用的蒸汽机实际上是2世纪亚历山大港的蒸汽动力玩具和17世纪德国真空泵的组合。钳工用台虎钳只不过是把两种古老的机械结合到了一起：杠杆手柄和螺纹装置。最早的自行车由三部分构成：轮胎、杠杆以及滑轮。除了创意，这些发明全都建立在统一的工业标准上。

如果说技术工业领域没有标准，那就有点言过其实了。技术发明依赖的通用标准不仅包括科学研究的自然规律，更重要的是测量方式的标准化，比如温度、质量以及电荷。不过大多数技术领域不像大自然，技术领域缺乏特定的标准化规范。自然界需要严格的标准化，因为它不像人类发明家，可以用额外的心力弥补工业标准上的不足。

功能不同的蛋白质，有的能催化反应，有的能推动分子转运，还有

的能维持细胞存活。这些功能的结构基础都相同，都是由氨基酸以同样的连接方式组合而成的。氨基酸之间的标准化连接方式是“肽键”，由一个氨基酸分子的氮原子和相邻氨基酸的碳原子构成。尽管每种氨基酸自身的结构不同，但由于“接口”的标准化，它们依旧能以相同的方式连接到一起。正是不同生物体内的氨基酸连接的标准化造就了我们熟悉的自然界。没有标准化，就没有超宇宙级数量的基因型。自然图书馆如果不能畅通无阻，生物进化也就寸步难行。

让组合优化成为现实的标准化不仅仅是蛋白质的专利，RNA也以标准的化学键连接单位分子。生命储存遗传信息的标准规范DNA使得细菌间的基因转移和性状组合成为可能。调控环路也以标准化方式调节着基因的表达，调节因子蛋白都能够识别和结合特定的DNA片段，通过改变不同基因前的调节片段，同样的调节因子能发挥不同的作用。我们手头只有一些为数不多的小部件，只要我们能够制订一种标准化的连接方式，然后以所有可能的方式对它们进行组合，无论这种组合多么盲目，我们创造新事物的潜力都已经和大自然不相上下了。

这种标准化的过程对于人类的工程技术领域来说显然是力所能及的：流行的乐高积木就是很好的例子，此外，另一项古老得多的技术也是很好的例证。

16世纪的威尼斯人安德里亚·帕拉迪诺（Andrea Palladino）是西方历史上最具影响力的建筑师之一。在漫长而光辉的职业生涯里，他至少为16个威尼斯最富有的家族设计过豪宅，修建过30座乡间别墅，还设计过数座教堂。帕拉迪诺的平面设计图总是风格迥异，每张设计图之间有着天差地别。他设计的建筑在建筑面积、户型、朝向以及房间布局上都不尽相同。

即便如此，这些房子依旧带有帕拉迪诺独特的风格烙印，虽然多数人可能说不清这种烙印究竟是什么。20多年前，艺术史学家乔治·赫西（George Hersey）与计算机专家理查德·弗里德曼（Richard Freedman）试图通过合作研究解释这种独特的风格，寻找隐藏在帕拉迪诺平面设计图中的秘密。他们的想法是，如果帕拉迪诺的设计图背后当真有特定的规律，那么我们肯定能用相应的算法模拟一张帕拉迪诺风格的设计图。

为了提取出帕拉迪诺风格的必要元素，赫西和弗里德曼分析了数十座帕拉迪诺式别墅的结构：房间的朝向、墙壁的布局、相邻房间的长度是否有特定的比例等。最后，他们成功了。两人设计的计算机程序最终

成功创造了一张帕拉迪诺风格的平面设计图。设计图里所有的细节都和已有的帕拉迪诺式建筑不同，无论是面积、朝向，还是房间布局，但是明眼人都能一眼认出这张新设计图里充斥着帕拉迪诺风格。

这个程序的算法在设计平面图时，会先勾出建筑的轮廓，这个轮廓往往呈矩形，然后算法会再用垂直或者水平的线贯穿矩形——这些直线代表墙壁，在建筑里分隔出不同的房间。一条直线可以把房子分隔成两个部分，而两条平行的直线则可以把房子分隔成三个部分，以此类推。每个房间继而以相同的方式被分隔，小房间再以这种方式继续分隔……直到房间的大小符合进一步设计的需要。对一个矩形用相互平行的垂直或水平线进行多次分割，每次用一条或多条直线，我们几乎可以绘制出无限多样的平面设计图。

帕拉迪诺的设计风格既不是刻意做作也不是随意而为，也没有听起来那么复杂高深。举例来说，如果一个房间被一条直线分隔成两个小房间，通常两个小房间的面积相同，或者其中一个是另一个的两倍大。而如果这个房间被两条平行线分割，通常中间的房间是两侧房间的两倍大。只要掌握这两点和其他几条规律并灵活运用，就算是计算机也可以模拟出这种最富盛名的文艺复兴时期的建筑风格。

自然界刻板的组合方式当然和计算机模拟帕拉迪诺风格的过程不完全相同。蛋白质是由更小的氨基酸组合而成的产物，而帕拉迪诺式建筑则是分割矩形建筑的结果。不过两者的共同点更值得关注：不管是蛋白质还是建筑学，都是用有限的基本元素和更有限的组织原则创造出种类庞杂的新产物。如果这个规律在工业革命前就已经存在于建筑学中，那么我们有理由推测，它也极有可能存在于工业革命之后的工程技术领域。

YaMoR就是一个很好的例子，它代表的领域正是现代数字技术。

控制机器人的电脑芯片上嵌有无数个晶体管，集成电路上的晶体管是一个只能对电压有无做出两种反应的简单电子元件：“0”代表没有电压，关闭开关；“1”则代表有电压，打开开关。所有的晶体管协同合作，对一串输入的数据流进行处理之后，再以一连串“0”和“1”的形式输出。每个晶体管相当于一个位，或者也叫二进制位。晶体管只有开和关两种状态，它们是计算机理解所有信息的基础。数学家对计算机工作原理的描述更精确，他们将计算机的工作状态描述为函数值的计算，即电路获得一个输入，经过演算和处理获得一个输出。电路演算所使用的函数，其名称出自英国数学家兼哲学家乔治·布尔（George Boole）1854年

的著作《思维规律的研究》。布尔创立的学术分支是科学上一个巨大的进步，而世人口中的布尔函数，早已经成为现代计算机科学的核心。

布尔函数中最简单的逻辑之一是与（AND）函数，你的每次搜索几乎都离不开这个函数。举个例子，如果你想找某首乐曲的电子版散页乐谱，譬如莫扎特的《魔笛》（*Magic Flute*），搜索引擎会检索所有标题中含有“莫扎特”的曲谱，对于每个标题，返还的结果要么为“是”，要么为“否”，用一个位来表示也就是“1”或者“0”。接着，引擎对“魔”字也会进行相同的处理。于是，就每个标题而言，两个关键词分别有两种返还结果，组合之后一共有4种不同的可能情况。数字编码的0-1阵列可以用数学家们习惯的方式书写，来替代布尔函数的函数值，这种写法被称为真值表。与函数的真值表写法如图7-1中a表所示，对于函数的每个输入，两个关键的检索结果分别在表格的左侧沿着纵向标出，表格右侧则是4个最终的返还值，也就是函数的输出，还是用“1”和“0”的形式标记。最终得到的4种结果中只有一种返还值为“是”——只有当先前两个输出的结果均为“是”，才算作《魔笛》检索的一个可能值。

如果你想要找标题中有“莫扎特”或者“魔”（或者兼有两者）的曲谱，搜索引擎需要运行另一个布尔函数：或（OR）函数。或函数和与函数的数据输入过程相同，都是检索所有含有“莫扎特”和“魔”的标题，但是它们判定的规则却不同。在或的逻辑里，只要输入数据的两个结果有一个为“是”，那么最终的输出结果就为“是”（如图7-1b）。所以，或函数的输出中不仅会有《魔笛》，还将有莫扎特的其他626支曲目，此外还有桑塔纳（Santana）的《黑魔女》（*Black Magic Woman*），史蒂维·旺达（Stevie Wonder）的《如果这就是魔法》（*If It's Magic*）等。布尔函数中还有一个更简单的逻辑函数，非（NOT）函数（如图7-1c）。非函数将输出所有返还值为“否”的结果，在这里，它可以帮你找到所有标题中不包含关键词“莫扎特”的曲谱。

a

莫扎特	魔	莫扎特与魔
1	1	1（是）
1	0	0（非）
0	1	0（非）
0	0	0（非）

b

莫扎特	魔	莫扎特与魔
1	1	1（是）
1	0	1（是）
0	1	1（是）
0	0	0（非）

c

莫扎特	非莫扎特
1	0（非）
0	1（是）

图7-1　真值表

与、或、非以及很多其他独特的布尔函数，比如XOR、XNOR、NAND以及NOR，帮助我们把自然语言中复杂的问题翻译成一串计算机能够理解的二进制数字。不仅如此，二进制数字与十进制一样，能够进行加减乘除运算。无论一台计算机有多么高端复杂，它的集成电路都在执行最基本的算数运算和简单的布尔函数，比如与函数。只需要两个最简单的数字，0和1，加上布尔函数，数字计算机就能够识别图片、对数据进行加密、发送语音邮件或是预测下周二的天气。如此看来，算数存在的意义远远不止是小学的数学考试而已。

布尔函数另一个非凡的特征是简单函数能够通过叠加组成复杂函数，在叠加的过程中，一个函数的输出可以作为下一个函数的输入。这就像一个乘法运算（4×3）可以用一个加法运算来代替：（4+4+4）。不仅如此，虽然理论上可以有无数种布尔函数，但每一种布尔函数都不过是与、或和非三个简单函数组合叠加的结果。这对于计算机来说非常重要，因为在集成电路中，晶体管往往通过串联形成计算单元以执行不同的布尔逻辑函数，这些晶体管单元因此被称为逻辑门。

图7-2中展示了几种简单的逻辑门，芯片设计师用这些示意图来代表与、或和非逻辑门。每个逻辑门的左侧都有一条或者两条直线代表输入，对应一个或者两个位。而右侧的一条直线则代表唯一的输出。图7-3中数个门电路被连接到一起以完成最简单的算数：计算两个二进制数字

的相加之和——如此简单的运算已经需要6个逻辑门，而每一个逻辑门中包含了多个晶体管。现代计算机能够加减乘除的极限当然已经远远不止64位，涉及的逻辑门数量也往往达到了百万级。

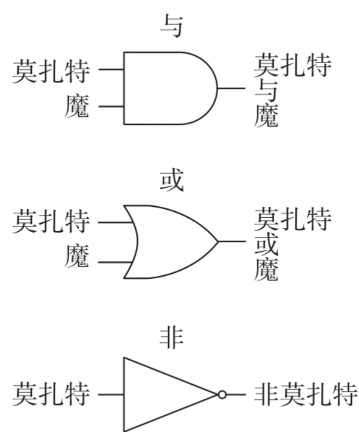


图7-2　逻辑门

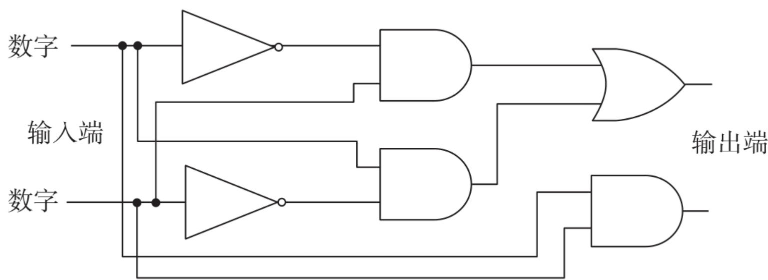


图7-3　计算二进制加法的电路

大多数集成电路在出厂前就会完成硬件连接，而像YaMoR这样的机器人则配备了可编程硬件，它们的芯片中某些逻辑门电路能够被修改，例如把某个与门电路改成或门电路。此外，不同逻辑门之间的组合方式也可以发生改变。有些可编程芯片甚至能在进行运算的同时修改逻辑门。逻辑门数量达到百万级别的电脑芯片已经不是小孩子手里的玩具了，而是灵活强大的计算引擎，它能帮助计算机学习许多人类才知道的东西。通过对自身硬件的修改，自主机器人不仅能移动，还能学会避开低洼的坑洞和其他陷阱。

如果你觉得上面这些听起来很熟悉，那是因为这与生物进化中一次改变一个分子的过程十分类似。可编程的逻辑门电路相当于可变的基因型，而不同的算法则相当于不同的表现型。和进化类似，计算机学习的

过程需要不断试错。在这个过程中，良好的表现会受到激励和加强，而不好的行为则会受到惩罚和削弱。当然机器人受到的惩罚往往不会像进化那么严厉。如果未来某一天你拥有的某个机器人高尔夫打得不太好，它大可以多多练习它的站姿、握杆或者甩杆技巧，而不是直接被淘汰。

另外，这种学习方式也不需要遗弃原有的知识。比如在学习高尔夫的前后，即便你脑中与坐、走、跑、跳等动作有关的神经环路逐渐发生改变，你依旧能够执行这些动作。逻辑门与生物进化的共同点还不止于此：逻辑门电路之间的连接属于通用连接，因为逻辑门的输出可以被任何其他逻辑门的输入所识别，就像蛋白质中标准化的肽键。只是对于蛋白质而言，肽键的合成、断裂和修饰要简单得多，而生产一块能够随意修改的标准化电路则要经过精心设计且耗费大量的人力。

标准化连接和少数几个基本逻辑门这两个条件已经足以打造出一款能够击败人类国际象棋冠军、从数百万页书中找出特定的一页或者“打印”3D物体的电脑芯片了。现实生活中的可编程芯片会让人联想到大自然的进化能力，如果有一座数字电路图书馆，里面收录了逻辑门电路所有可能的组合方式，那么它的组织形式会不会和自然图书馆一样？这个问题的答案将告诉我们，生物进化的曲速引擎是否有适用于工程技术发展的改装版本。

卡蒂克·拉曼（Karthik　Raman）为我们找到了答案。拉曼毕业于印度最顶尖的大学之一印度理工学院，他选择了我的实验室继续博士后研究。而拉曼可不是空手而来的，他带着他对于科学的极度狂热，面对失败毫不气馁的毅力以及分析复杂数据的鬼才天赋，跨进了我的实验室。当我提出让他研究可编程电路图书馆时，他二话不说就扑到了这项研究上。

虽然如今市面上的可编程芯片中，逻辑门的数量往往在百万级。但是经过严谨的估算后，我们认为研究规模相对更小的集成电路是一个更好的选择。如果我们考量一个包含16个逻辑门的集成电路，可能的电路数量将达到10^{46}个，这个数字会随着逻辑门的增加以指数级增长。当逻辑门达到36个时，电路的可能数量已经超过了10^{100}个。巨大的基数倒是让是否要制作芯片进行实物测试的疑问显得清晰明了：面对数百万个需要测试的电路，我们也只能用计算机对它们进行模拟了。

包含16个逻辑门的集成电路理论上能够演算10^{19}种布尔函数，不过我们以前并不清楚图书馆内的电路是否都有功能。说不定其中大部分的电路只能执行一些比较低级的计算，比如加法和乘法。为了寻找答案，

拉曼首先构建了一张巨大的网络，并向其中添加尽可能多的集成电路。他设计了200万种集成电路，每种电路中的逻辑门之间都以随机的方式互相连接。这200万种不同的电路可以演算出150万种不同的功能函数，而其中只有寥寥数种函数类似于我们熟悉的与函数。虽然拉曼在研究中已经尽了很大努力，但是他的工作仍然只涉及了一小部分集成电路，有待验证的集成电路数量是200万的10^{40}倍，而布尔函数的数量则是150万种的10^{12}倍。拉曼的工作告诉我们，即便是简单的集成电路，同样能够计算数量庞大的布尔函数。

由于图书馆中电路的数量要远远多于布尔函数，确切的数字是10^{26}倍，所以图书馆中许多电路表达的含义应当是相同的，含义相同的文本执行相同的逻辑函数，但是我们并不知道它们的组织方式。为此拉曼随机选定了一个执行任意函数功能的集成电路并寻找它的相邻电路。他寻找的方式是对集成电路中的逻辑门进行调整，例如把某个逻辑门的输入端调换到另一个逻辑门的输入端。

如果这种“突变”后的集成电路依旧可以执行原来的逻辑函数功能，拉曼就选择留下它。如果逻辑函数改变，他就重新尝试其他连接方式。通过重复这个过程，拉曼的集成电路一步一步地远离最初的起点，而电路的函数功能保持不变。拉曼从某个随机集成电路出发，在保证函数功能不变的情况下进行随机游走，重复了上千次。

集成电路网络中随机游走能够到达的距离，甚至比我们在之前章节中探讨过的基因型网络都远：大多数集成电路可以在保证函数功能不变的情况下，从图书馆的一端走到截然相对的另一端。两个集成电路除了执行的函数功能相同之外，可以说毫无相似之处，从单个逻辑门到多个逻辑门之间的连接方式皆不同，但是它们的确都位于同一张集成电路网络中，只要不断改变基本逻辑门之间的连接就可以把其中一个变成另一个。不仅如此，在研究中我们还发现，无一例外，所有的函数都具备这个性质。也许这是所有二进制逻辑门电路的一个基本特性。数字图书馆和生物图书馆相比，也许有过之而无不及。

下一步，拉曼把注意力转向了集成电路所在的社区，他首先寻找某个电路所在的社区，确保这些相邻的集成电路与原电路功能相同，然后再寻找社区电路的相邻电路，列出它们所有的函数功能。拉曼发现，这些社区的多样程度与生物图书馆不分伯仲。功能相同的不同集成电路所在的社区中，有超过80%的集成电路执行着不同的函数功能。和生物图书馆类似，对于集成电路而言这是一件好事：一个集成电路能够维持自

身的函数功能，同时又保留有巨大的改变潜力。虽然一个集成电路的所有相邻电路仅有大约60种新的函数功能，但是只要对逻辑门的连接进行10次改变，潜在的新函数数量就达到了100多种；100次改变后，这个数字变成了400多种；而如果我们从起点走出1 000步，那么沿途将遇到近2 000种新的函数。

在先前的章节中我还打过多维空间网络的比方，基因型网络就像一块复杂得难以想象的编织物，而且这块织物只存在于高维空间。拉曼发现这个比喻同样适用于集成电路，功能特定的某个集成电路能够以任何一个随机电路作为起点，通过改变若干个逻辑门之间的连接而获得。数字电路也像一块高维空间的纺织物，这张网络让我们寻找所需功能函数的效率得到了大大提升。

通过与基因型网络的对比可以得知，集成电路网络具备推动电脑芯片优化的所有特征，它就是硬件进化的曲速引擎。未来的某一天，YaMoR的继承者将不仅能够经过学习避开路上的坑洞，还可以学会更复杂的技能，比如洗碗、照顾孩子打球等。它们的数字大脑可以通过按部就班的修改和优化，在不影响原有行为的基础上习得新的技能，温故而知新。如果有人说我们的大脑也在用相同的方式进行学习，我一点都不会感到惊奇。如今我们都知道，在人的一生中，大脑中神经元之间的突触连接始终在发生着变化，而这种变化与生物探索基因型网络的方式相似。倘若如此，工程学借鉴生物进化的日子将指日可待。

不幸的是，我们对工程学该如何借鉴生物学依旧一无所知。我们仍然不知道创造力的物质基础到底是什么。不过我们也发现，新的技术发明不是免费的，因为拉曼找到了标价签，而这个代价我们并不陌生。

拉曼对逻辑电路的复杂性进行了分析，也就是电路中包含的逻辑门的数量多少。他找到了最简单的逻辑电路，所谓最简单，就是无论多小的改变都无法保证电路的功能不变。最简电路中的每一个逻辑门、逻辑门之间的每一处连接都至关重要，稍作改变，电路的函数功能就无法维持。因此，最简电路的结构和算法无法被改进，经得起修改和优化的电路需要一定的复杂度。越复杂的逻辑电路对修改的耐受性越高。

复杂电路中看似多余的逻辑门和逻辑门连接就像备用零件，用以帮助芯片学习新的函数功能，它们就是爱迪生所说的“无用的垃圾”。和生物学中一样，看似多余的复杂性，其实对进化而言至关重要。这就是人类的工程技术能够从自然界借鉴的东西之一：如果我们想撬开创造力的黑匣子，奥卡姆剃刀的刀刃多少显得有些单薄。如果说精简主义是水，

那么创造力就是油，两者互不相融。

但是这并不意味着技术创新的领域里容不下精简主义和高雅主义。恰恰相反，只是它们藏在更深的层次里。高雅主义的实质就是精简主义本身：利用有限的原料和有限的规则，创造出世界万物。利用这个规则，大自然创造了蛋白质，创造了调控环路，创造了新陈代谢，创造了生命。从简单的病毒到复杂的人类，继而催生了我们的文化和技术：无论是《伊利亚特》还是iPad。技术发明的精简主义和高雅主义就像自然图书馆，隐藏在现实世界的背后，看不见，摸不着。我们只能从生命之树上觅得一些亦真亦幻的风影，就像柏拉图洞穴里的变幻之影。

柏拉图的洞穴

1970年10月，科普杂志《科学美国人》（*Scientific American*）上刊登了一篇文章，报道了英国数学家约翰·康韦（John Conway）设计的一款生命游戏“Game of Life”。这个“游戏”的参与者不是人类，而是计算机中以小方格代表的细胞，每个细胞有“开”（生）和“关”（死）两种不同的状态。在计算机内的网格图中，每个小方块被另外8个方块包围。游戏的规则很简单：如果每个细胞周围的细胞中存活的细胞少于2个，那么它的状态就转变为“关闭”，按照游戏中的术语来说，就是“它死了”。如果细胞周围存活细胞的数目大于或者等于4个，它也会死去。而如果细胞周围活细胞的数量是2个或者3个，它就能够存活。最后一条游戏规则是：如果一个死细胞周围有且仅有3个活细胞，那么它将复活。这个游戏的设计原型是如何构建一台自我复制的机器，它的提出者是一代通才约翰·冯·诺依曼（John von Neumann）。

这就是生命游戏的全部规则。游戏一旦开始，所有细胞的生死存亡都以上述规则为准，而游戏接下来发生的一切远远没有人们想的那么简单。计算机中的图形变化模式相当复杂，变化样式无穷无尽，无法预测，例如有的细胞可以通过自我复制完成“传宗接代”。从简单的开局开始，生命游戏可以无限进行下去，不断生成复杂的图形，变幻莫测。

这简直就是真正的生命！

生命游戏更像是一则生命的隐喻，而不是对生命的模拟。它带给我们的启示非同寻常：用数学和计算机科学研究生命的多样性是可行的。当然这个想法在生命游戏出现之前早已萌芽。在《物种起源》出版17年之后，达尔文在他的自传中写道：“我一直懊悔于没能在数学方面有所精进，领略它的博大精深。我羡慕那些颇有数学天赋的人，他们似乎有着异于常人的洞察力。”

达尔文逝世4年之后，萨莉·加德纳发明了诡盘投影仪，他们两人的工作都在各自的领域中掀起了革命。不过，就算达尔文真的是一位数学家，他依旧会迷失在寻找生命建筑师的黑暗里，或者更糟，他可能根本就不会意识到进化的存在。在巨大的自然图书馆里，光有达尔文的进化论还远远不够，生物进化的科学理论之光需要更多燃料。

生物学和数学需要更立体的交织，当这一点实现时，时间已经又过去了将近一个世纪。休厄尔·赖特和费希尔是这个领域的先驱，他们用数学理论填补了传统达尔文进化论和孟德尔遗传学之间的空隙，首次通过计算预测了自然选择对推动生物进化的助益程度，为现代进化论的建立奠定了基础。而后半个世纪转瞬即逝，直到系统生物学出现，我们才窥见微观分子的相互作用如何造就了生物复杂的行为和表现型。当然，如果不是这些研究，我们也不会知道细胞比生命游戏中的方块要复杂得多。细胞的调控环路与人类大脑中的神经网络类似，它们都能通过执行复杂的计算以调节分子行为，维持自身的存活。虽然生物的调控环路与数字计算机差异巨大，但是拥有实体的生物与虚拟的数学和算法之间却有着某种深深的羁绊，这连康韦和达尔文都没有猜到。

系统生物学中强大的数学工具让我们能够破译自然图书馆中每个基因型所对应的表现型，这是我们利用计算机研究生物进化的先决条件。在研究的过程中，我们意识到了基因型网络的存在，而基因型网络正是新性状的源头（不管是新的代谢、生物大分子还是调控环路），也就是生命的源头。生命从混沌之初发展到单细胞生物，从细菌、真菌的远古祖先到原始的蠕虫、鱼类、两栖类、爬行类，直到人类出现，经历了万亿年的繁衍生息。而在这个过程中，基因型网络功不可没。

光有基因型网络还不够，如同宇宙尘埃只是循着引力的引导就形成了巨大的星系，数学让我们意识到自然图书馆在自组织过程中也遵循一个简单的规律。这个规律就是普遍存在的发育稳态，而遗传上的稳定性来自复杂性。一定程度的复杂性让生物能够适应不断变化的环境。由此，我们发现了自然图书馆中馆藏之间紧密交织的相互关系。

我们在书中探讨的图书馆与解剖学家研究的肌肉、神经和结缔组织完全不一样。它们甚至不像单细胞生物那样能够用显微镜看到，也不能用观察DNA晶体结构的X射线成像。它们是触不可及的抽象概念和数学模型，只允许我们借助思维徜徉其中。

那这些图书馆是不是只存在于我们的想象中呢？它们是客观存在的，还是我们凭空捏造的呢？

那么知识（尤其是数学）到底是客观存在的，还是我们发明创造的呢？哲学家们对这个问题的讨论至少可以追溯到毕达哥拉斯或者柏拉图，已经有超过2 500年的历史了。柏拉图认为，人类肉眼可见的世界不过是墙上的虚幻投影，在其之上是更高阶的理性世界，正是这个更真实的世界将自己的影子投射在感官世界的墙面上。柏拉图学派认为，客

观的真理是存在的，只是它们存在于更高阶的理性世界之中。真理本身并不以人类的存在为前提，比如不管人类有没有观察或者意识到，月球的背面一直都客观存在。而相反的声音也一直存在，比如奥地利哲学家路德维希·维特根斯坦（Ludwig Wittgenstein）就认为，数学是人类发明的产物，用维特根斯坦自己的话说就是“数学家是发明家，不是发现家”。

柏拉图学派在这场辩论中略占上风，虽然最好的论据可能连柏拉图本人都不知道。这个论据就是数学理论与物理事实之间惊人的吻合性，借用伽利略的一句名言：“数学是上帝书写宇宙时用的语言。”所有天真的创世论者都应当听听这句话。诺贝尔奖得主、匈牙利籍物理和数学家尤金·魏格纳（Eugene Wigner）曾经说：“数学在研究自然科学中简直有用得难以置信。”

确实难以置信。我们不知道为什么牛顿定律除了预测苹果下落的速度之外，还能用来计算行星的自转、预测星系的形成。但它们就是可以。此外还有无数数学定律，它们能够解释许多时空上遥不可及的现象，我们甚至都不用直接观察或是体验到它们。数学与现实之间的联系着实紧密，瑞典理论物理学家迈克斯·泰格马克（Max Tegmark）[27]甚至直言，宇宙的本质就是数学。

但如果仅仅是“有用得难以置信”，还不足以证明自然图书馆和基因型网络的真实性。让它们进入我们视野的另一位功臣是21世纪的工程技术。有了技术的帮助，一场纸上谈兵的辩论转而变成了一门实验科学。转变发生的原因是我们拥有了读取自然图书馆中馆藏的能力。比如，我们可以构建蛋白质图书馆中的任何一本馆藏，然后用生物化学手段解读它的化学含义。自然界的生物早就先我们一步找到了不少蛋白质，而它们在生物体内的作用常常出乎我们的意料，比如抗冻蛋白、晶体蛋白和Hox调节因子家族。我们几乎可以肯定，自然图书馆中还有无数的惊喜，远非人类的发明创造所能企及。

当我们探讨自然图书馆时，我们的目的不仅仅是寻找生命的起源与科技创新的来源。我们希望哲学中最历久弥新的命题能够引发读者的些许思考。我们的研究发现，推动生命进化的动力可能比生命本身更古老，甚至比时间更古老。



扫码获取“湛庐阅读”APP，
搜索“适者降临”查看本书参考文献。

听闻此书即将交付印刷出版，内心百感交集。

与此书的相遇缘起于大学时的同学侯新智，侯作为科班的语言专业出身，领我入了英文书籍翻译的门，只可惜我到现在为止依旧是半路出家的水平。在湛庐文化的第一次翻译经历始于侯先斩后奏将手头正在翻译的半本书交给了我，在此也感谢简学老师对此的宽容大度。第一本书完成之后，简老师一错再错，把这本《适者降临》送到了我手上。非常感谢两位提供的机缘。

郝莹老师现在步了简老师的“后尘”，承蒙错爱。

翻译的过程说不上艰难险阻，但也并非顺风顺水。对于一个哲学门外汉来说，书中大量的哲学内容探讨让我甚是头疼，每每碰到理解上的难题时，幸而有哲学系的同学赵梦雪帮我出谋划策。书籍翻译不在势头猛，而在细水长流，翻译之中曾经一度陷入过困顿，这里也特别感谢同学赵璐洁的帮助。以上两位现在都在欧洲求学，遥祝她们一切顺利，学有所成。

能够坐在笔记本前，敲下漫漫十多万字，少不了其他人在生活上的帮助。在这里，要感谢杨传剑，感谢室友陶昶煜、李文睿和朱奕彰，感谢同学丰俊林、杨子逸、杨樾、王国乾，感谢郑棒、李华铭、刑美英、吴宇佳，感谢老许。除此之外，承蒙太多关照，无法在这里一一赘述。

翻译过程中恰逢自己毕业离校，难免面对熟悉的人各奔前程。如果说毕业典礼上的风光犹如浓缩了昔日在校期间的所有精华，那么平平淡淡的校园生活就相当于在这场浓缩里扮演了跬步和小流的角色。前者是突变论者喜闻乐见的场景，而后者则是渐进主义者青睐的论调。初入社会的不少同学在面对毕业这个突变时显示出了前所未有的谨慎，因为这意味着之前平静和按部就班的生活随即荡然无存。

也许人们对于生活的期许有赖于对生活本身的看法，尤其是对于生活到底是离散还是连续的看法。相信生活是离散的人大概会更看重生活中的突变，就像相信生物进化依赖于突变的孟德尔学派；认为生活是连续的人大概更愿意强调眼前的每一件小事，正如愿意在生物进化中寻找每个微小改变的渐进主义学派。不过正如书中所说，突变论和渐进论的争论将生物进化的问题带上了一条歧途。问题的关键不在于变化剧烈与否，而在于给生活带来变化的动力起源于何处。

早先有许多相识的人会诧异于我在翻译文章和书的事实。我的专业和翻译没有任何关系，甚至于可能几乎不用跟英语打交道。外语和语文本身并非我的兴趣，后者几乎可以说是软肋。也许大家常常会太过于关注一个突变的实际意义，突变可以是单基因突变，乃至于单核苷酸突变。除了错义突变和无义突变之外，单核苷酸突变也可以是同义突变。这类似于日本科学家木村资生提出的中性突变理论的事实基础：大部分的突变都缺乏进化上的意义。它们既不好，也不坏。

我们起于殊途，也许归于同处。

中性突变出的翻译功底，想必会有不少文字和内容上的纰漏，望指正，望海涵。

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读APP：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当做阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”APP。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读APP玩转指南

湛庐阅读APP结构图：



三步玩转湛庐阅读APP：



APP获取方式：
安卓用户前往各大应用市场、苹果用户前往APP Store
直接下载"湛庐阅读"APP，与最聪明的人共同进化！

使用APP扫一扫功能，
遇见书里书外更大的世界！

快速了解本书内容，
湛庐千册图书一键购买！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！



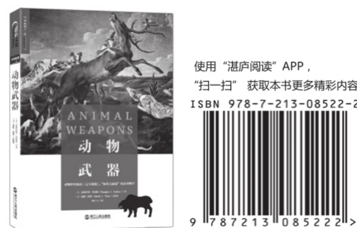
你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

湛庐文化
Cheers Publishing
a mindstyle business 与思想有关

延伸阅读

《动物武器》

- ◎ 动物世界版的《竞争战略》，“知识大融通”的新博物学。从动物武器到人类战争，一本书读懂武器进化史。
- ◎ 著名学者、“伯凡时间”创始人吴伯凡，华大基因CEO尹烨，北京大学生命科学学院教授谢灿，《三联生活周刊》主笔贝小戎，国家博物馆讲解员河森堡，社会生物学之父爱德华·威尔逊等知名大咖鼎力推荐！
- ◎ 荣获美国大学优等生联谊会“科学图书奖”，一本有趣又有料的“新博物学”佳作！



《半个地球》

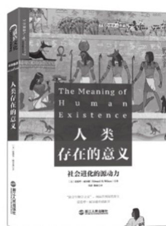
- ◎ “社会生物学之父”、两届普利策奖得主、进化生物学先驱、殿堂级的科学巨星爱德华·威尔逊重磅新书！
- ◎ 北京大学哲学系教授刘华杰，中国科学院大学教授李大光，分享收获农场执行董事、“全球40岁以下影响食物系统的20人”农人石嫣，《大转向》作者史蒂芬·格林布拉特，科普作家奥利弗·萨克斯，著名全球发展问题专家杰弗里·萨克斯鼎力推荐！
- ◎ 爱德华·威尔逊继《生命的未来》与《缤纷的生命》之后又一聚焦生物多样性、关注全球物种灭绝的倾情力作！亚马逊年度最佳科学图书！





《人类存在的意义》

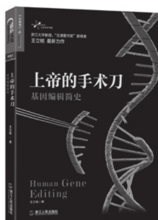
- ◎ “社会生物学之父”、两届普利策奖得主、进化生物学先驱、殿堂级的科学巨星爱德华·威尔逊最新力作！
- ◎ 北京大学哲学系教授刘华杰，美国前副总统阿尔·戈尔，环境保护主义理论家、畅销书《幸福经济》作者比尔·麦吉本，著名脑神经学家、科普作家奥利弗·萨克斯，著名全球发展问题专家、畅销书《贫穷的终结》作者杰弗里·萨克斯鼎力推荐！





《上帝的手术刀》

- ◎ 雨果奖得主郝景芳、清华大学教授颜宁倾情作序！雨果奖得主刘慈欣、北京大学教授魏文胜、碳云智能首席科学家李英睿、《癌症·真相》作者菠萝、《八卦医学史》作者烧伤超人阿宝联袂推荐！
- ◎ 一本细致讲解生物学热门进展的科普力作，一本解读人类未来发展趋势的精妙“小说”。







[1] 诡盘投影机，一种将一系列静止照片滚动播放形成运动画面的仪器。——译者注

[2] 波兰天文学家哥白尼在1543年的《天体运行论》里提出了“日心说”。但是早在古希腊，阿那克西曼德就已经提出了“太阳中心”的宇宙观。——译者注

[3] 用进废退，指经常被使用的生物器官和功能得以增强，而不被使用的则会退化。——译者注

[4] “DNA之父”，20世纪分子生物学的带头人之一。其讲述DNA双螺旋结构发现历程的著作《双螺旋（插图注释本）》中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

[5] 美国生物学家，被很多人称为生物学界的“坏小子”，讲述其在生命科学领域做出的重大研究的著作《生命的未来》中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

[6] 退火：在生命科学领域，退火指降低反应温度，让变性的核酸链复性，重新形成双链的过程。——译者注

[7] 误差灾变论，指生物体在复制遗传物质时积累大量的突变而导致的物种灭绝模型。——译者注

[8] 又叫柠檬酸循环，是需氧生物体内普遍存在的代谢途径，分布在线粒体中。——译者注

[9] 活力，指让生命体区别于非生命体的神秘力量，不受自然法则支配。——译者注

[10] 又称蓝藻细菌，其尸体在距今35亿年前遗留的钙质形成了现今的叠藻岩（stromatolites）。

[11] 五氯苯酚（Pentachlorophenol）的属名Sphingobium意指其为鞘脂菌属，种名Chloroplenolicum暗示其可以利用和吸收五氯苯酚。——译者注

[12] 齐克隆B（Zyklon B），德国化学家弗里茨·哈伯（Fritz Haber）发明的氰化物药剂，原为除虱的杀虫剂，后用于奥斯维辛集中营，屠杀了大量犹太人。——译者注

[13] 其实这里作者的描述并不准确，1/2并不能代表一种生物中一半的酶可以被另一种生物合成。——译者注

[14] 杜威十进制图书分类法，由美国图书馆专家麦尔威·杜威（Melvil Dewey）发明，为多数英语图书馆所沿用。——译者注

[15] 美国国会图书馆分类法，是指在美国国会图书馆馆长普特南的主持下，根据本馆藏书编制的综合性等级列举式分类法。——译者注

[16] 这个名字源于阿拉伯树胶，后者是合欢树合成的天然树胶。

[17] 通常，美国高等院校中教师每7年会有一次年假。——译者注

[18] 这段为字母乱码，无实际含义。——译者注

[19] 《第二十二条军规》，美国作家约瑟夫·海勒的讽刺小说，故事主线围绕一条自相矛盾的军规展开。——译者注

[20]　这里所说的折射角是指水中的光线与水面垂直线所成的角度。——译者注

[21]　雅典娜的诞生：宙斯吞食墨提斯之后患上了头疼，众神赶到帮助宙斯打开脑袋之后，里面蹦出了身披铠甲的雅典娜，雅典娜因此被奉为智慧女神。——译者注

[22]　腻虫和蜜虫为中文俗称，原文为blackfly，greenfly。因未能查证其确切的中文翻译，故粗译为此名。——译者注

[23]　爱尔兰大饥荒始于1845年，终于1952年，造成饥荒的原因涉及自然、政治和经济等多个方面，是历史学家研究爱尔兰的重点事件。——译者注

[24]　19世纪中期发生于法国的大规模蚜虫虫害，对许多葡萄酒庄园造成了毁灭性打击，其后多年法国葡萄酒行业一蹶不振。——译者注

[25]　轻度肺炎（walking　pneumonia），“可以走路的肺炎”，原意为不需要入院或者卧床休息，俗称轻度肺炎。——译者注

[26]　《技术的本质》中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

[27]　平行宇宙理论世界级研究权威、MIT物理系终身教授。其权威巨著《穿越平行宇宙》中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

CHEERS
湛庐

人类的价值

合作如何让我们
成为强大的群体
How Culture Transformed
Our Species


圣塔菲书系 Santa Fe Institute Series

A Different Kind of Animal

[美] 罗伯特·博伊德 著
Robert Boyd
袁冬华 译

模仿学习，
让人类更好地适应环境

社会规范，
让大规模合作成为可能

浙江教育出版社

版权信息

本书纸版由浙江教育出版社于2019年12月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：人类的价值

著者：罗伯特·博伊德

电子书定价：71.99元

A Different Kind of Animal: How Culture Transformed Our Species by Robert Boyd

Copyright © 2018 by Princeton University Press.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

是什么让人类如此独特

斯蒂芬·马塞多
普林斯顿大学人类价值研究中心主任

是什么让人类如此独特？是否像很多人所认为的那样，因为拥有更高级的智力，人类才得以与其他物种区分开来？

人类进化与社会变革领域的知名教授罗伯特·博伊德对这一命题进行了重新梳理，并驳斥了这种存在于大众之中的普遍观点。一开始，博伊德并没有谈及那个更为人熟知的问题，也就是人类为何如此独特，而是提出了另一个问题，即人类为什么能够取得生态学意义上的广义成功，从而超越其他物种，比如，人类能够适应地球上各种各样的栖息环境并繁衍至今。早在10 000年前，人类就已经占据了地球上的大部分地区，除了南极洲和几个偏远小岛之外，没有其他生物能够做到这一点。如果不是“超级大脑”，那又是什么让人类获得这绝无仅有的地位呢？

人类之所以能够适应各种不同的环境，并解决各种问题，主要依靠的不是个人智慧，而是“累积性的文化适应”（cumulative culture adaptation），以及更长时间维度下，不同文化之间的达尔文式的自然选择过程。这里的“文化”主要是指各种社会规范和道德价值。博伊德认为，不仅人类是自然世界的一部分，人类文化也是自然世界的一部分。文化使人类成了一种与众不同的生物，文化是人类生物学的组成部分之一，让人类形成了独特的骨盆，以及臼齿上厚厚的釉质。

通过与多人的合作，尤其是与人类学家彼得·里克森（Peter Richerson）的合作，博伊德得以在三十多年来一直专注于人类进化过程中的一个关键论题，也就是人类群体文化信息的传播机制。这种理论又被称为“双重继承”理论或“文化进化”理论，“文化群体选择”是其子理论之一。在接下来的各章中，博伊德对该研究项目的主要内容和观点进行了梳理、详述，并对其他评论人的评论做出了回应。“博伊德所建构的框架深刻地揭示了人类本性，对人类社会的运作机制有着深远的启示，社会科学和人类文化领域的学者应对这些研究成果给予更多关注。”这是经济学家保罗·西布莱特（Paul Seabright）在本书评论部分中所提出的观

点，我对此深表赞同。对于门外汉或略有所知者来说，本书是了解进化领域的极佳选择。

本书滥觞于一场以“人类的价值”为主题的坦纳讲座（Tanner Lectures on Human Values），该讲座是由普林斯顿大学人类价值研究中心于2016年4月举办的。

博伊德在论述时不仅旁征博引，还穿插了各种轶闻趣事，以期证明人类运用洞察能力和适应能力的方式是我们未曾知晓的。人类最常用的学习方式并不是了解事物的前因后果，而是模仿那些已掌握了有用的地方性“秘诀”的人。博伊德认为，选择会偏向这种心理机制，以致大多数人会因为周围人都信奉某个事物，而去信奉同个事物。事实上，他认为，即便是在最简单的狩猎-采集（hunter-gatherer）社会中，获得所需工具和知识过程的复杂程度也远远超出了个体能力。文化是累积性的、地方性的隐性知识的“储藏室”。累积性文化进化是人类最大的，也是最独特的优势。

结果并不总是积极正面的，非适应性的理念和错误的观点会通过盲目模仿而被传播。不过，正如保罗·西布莱特所认为的，就总体效果而言，极为灵活的合作能力，使人类变成了一个聪明的群体，而且聪明的程度超过任何个体。事实上，因为拥有模仿学习的能力，因此“个体所需达到的聪明程度，并不需要企及群体的聪明程度”。

在第2章中，博伊德利用社会学习和累积性文化进化的观点，进一步阐释了人类社会如何适应不断变化的环境，并开发出越来越复杂的工具和技术。凭借模仿学习的能力，以及进化而来的“信任心理”机制，便可以解释为何社会规范会是人类生活的核心，为什么是人类，以及人类是如何成为“超级合作者”的。即便是在狩猎社会，人类的合作程度也大大超过了其他物种。数千年以来的累积性文化进化帮助人类构建出了一个庞大的“世界性分工和交换活动网络”。人类的独特之处在于毫无关联的个体可以结合成庞大的群体，互相合作并提供公共产品。

因此，博伊德提出了这样一个论题：在自然界中，所能看到的大规模合作只存在于亲属之间，但人类社会并非如此。那么，自然选择是如何改变人类心理机制，从而激发毫无关联的个体结合成庞大的群体，并进行合作的呢？

要想进行大规模的合作，就需要构建行为规范系统，并通过制裁来强化这一系统。在规模更大、复杂程度更高的社会里，合作和公共产品

的供应非常依赖具有制裁功能的第三方，比如警察局和法院等机构。

与合作有关的社会规范，不仅形式多样，还进化出了多个立足于不同领域的道德准则，比如结婚和继承方面的准则，以及各种政治系统。不同的社会和文化群体在这些道德准则的基础上展开竞争，而这些道德准则在不断变化的环境中有着不同的传承力，在社会竞争中也有着不同的传播力和占据优势的力量。比如，罗马帝国之所以要放弃最初所信奉的宗教转而信奉基督教，是因为“原先的宗教并没有互助的传统”，而基督教社区则“照顾病患”，减少了死亡率，提升了人们的福祉。当然，罗马帝国的政治机构因适应了当时的环境而得以续存。

博伊德尖锐地指出，从进化的角度对社会生活所做的解释，往往会给人一种自由意志论的错觉，让人觉得可以把社会定义为“由自利性的个体和血亲家庭之间的裙带关系所构成的网络”。他驳斥了这种论断，坚持认为，即便是小规模的合作，也离不开共同信奉的行为规范的制约，而这种规范往往需要通过第三方制裁来强化。不仅如此，这些规范还为人们的持续性合作提供了重要的支撑。

在我看来，博伊德的整个学术研究以及这些演讲中最突出的特点是在研究取向上坚持多学科并举，在立论时旁征博引，且论述材料丰富翔实。他既能以文化人类学家的身份和学生们一起在多地开展田野工作，也能借助“理性选择”理论进行数学建模，还善于借鉴自然科学和社会科学领域的观点与方法。在第7章中，他谈及了与彼得·里克森的长期合作，并谦虚地回应：“我们的研究方式过去是，现在依然是广泛阅读人类学、心理学和经济学等领域的文献资料，挖掘有前瞻性的经验问题，而后用种群生物学理论来解决它们。”说起来简单，但很少有学者能够企及他所做研究的广度。

不要把“文化群体选择”和道德以及其他进化形式混为一谈。博伊德的理论属于社会科学范畴，是实证研究，而不是伦理规范性研究。当然，任何社会都离不开社会规范和广泛合作，因此，毫无疑问，对于那些关注“充满正义的美好社会应该建立在何种人性基础之上”“应该建立什么样的社会及社会制度基础”的人们来说，本书是非常值得一读的。

博伊德最后所做的总结强调了其核心观点，即“文化的适应性进化”是获得生态学意义上的成功，以及拥有合作能力的“基本要素”。人类之所以能成为万灵之长，是因为没有其他生物能够像人类这样，适应如此多样的地域环境。这种适应性，单靠个体的发明创造是无法做到的。

在博伊德的两个演讲之后，是四位来自不同学科领域的杰出学者所作的评论。

艾伦·奥尔（H. Allen Orr）是一位进化生物学家，他不仅研究物种的形成和适应性问题，还针对各类受众的需求进行写作。对于博伊德的文化学习模型，奥尔提出了两个比较重要且有趣的问题。第一个问题是，博伊德是否夸大了“超级大脑”模型与其提出的模拟模型之间的对立性。“超级大脑”模型强调的是从认知的角度对人类的成功进行解释。奥尔认为，要想成功地进行模拟，往往需要借助神经机制方面的重要支撑。此外，奥尔还指出，博伊德的观点与著名经济学家、社会理论家弗里德里希·哈耶克（Friedrich Hayek）的观点有契合之处。哈耶克也强调，社会的成功和进步离不开对隐性的、分散的地方性知识的运用，离不开对社会规范和伦理道德的文化传承，也离不开社会进化所产生的各种制度。第二个问题是，西方的自然科学家和社会科学家是否因为哈耶克的政治立场，而未能对其学术理论给予足够的关注。

金·斯特林（Kim Sterelny）是一位著述丰富的科学哲学家，专注于进化生物学领域的研究。他与其他几位评论人一样，对于博伊德的主要观点持赞同意见：人类对各种环境的适应能力，在所有生物中是独一无二的。但是，与奥尔和西布莱特一样，斯特林心里也存在着这样一个疑问：在解释人类对生态环境的超级适应能力时，博伊德是否对“人类独特的智力”不够重视。对于博伊德所批驳的“图书馆”模型和“超级大脑”模型，他持部分认可态度，并认为隐性的实践性诀窍的确是一种知识。此外，斯特林还认为，博伊德简化了社会学习，把社会学习看成“依照惯例的”，或者说是“信任背景下的”启发式学习法，并且过于依赖这种简化。斯特林质疑的最后一点是：在不同的社会和时间背景下，社会学习是否会发生变化，以及如果会发生变化将如何发生。

鲁思·梅斯（Ruth Mace）与博伊德一样，是一位人类学家，虽然她的学术取向是人类进化生态学，但她还是对博伊德在研究人类进化时所坚持的多学科取向理念称赞有加。当然，她在评论中还强调了实证检验的重要性。她认为，依然有许多问题需要回答，比如规范是如何出现的，为什么会发生变化，这些规范是如何得到强化的，变革的难易程度如何，等等。她的批评性意见更多一些。她还指出，博伊德认为一些行为是由社会规范和制裁所激发的，但实际上，可能是由个体自利性动机所激发的，比如，“决定参加战斗”之类的行为。她描述了自己在北爱尔兰开展的关于群体冲突的实证研究，并提出了这样一个问题：群体之间

的竞争和冲突，譬如种族之间的战斗，是否会导致狭隘的利他主义？

保罗·西布莱特是一位涉猎广泛且颇具影响力的经济学家。他认为博伊德的研究忽视了一个方面，也就是那些“形成人类阴暗面的因素”。西布莱特认为，如果人类既是这个星球上最具生态环境适应性的物种，也是合作规模最大的物种，那么人与人之间的竞争，在残酷和暴烈程度上也是所有物种中最突出的；操控同类的狡猾本领亦是所有物种中最厉害的。西布莱特评论说：“在大多数社会中，大部分围绕规范所进行的沟通都会涉及某些个体对另一些个体的操控。”但这种操控在博伊德所列举的例子中却几乎看不到。

在本书的最后，博伊德做出了回应。他认为四位评论人虽然各自提出了一些反对意见，但都是经过深思熟虑的，由此深表感谢，毕竟这些评论都认可文化进化是非常有价值的。

博伊德注意到，鲁思·梅斯、金·斯特林和保罗·西布莱特都对他不太相信人类能够广泛搜集信息并做出明智决策这一观点提出了质疑。对此，博伊德坚持己见，认为虽然个体的选择确实能发挥一定的作用，但人们的基本观念却来自所处的社会。

就艾伦·奥尔等人的相关评论来说，基本上可以认定，博伊德赞同这个看法，也就是认知能力和文化学习是互相强化的。

面对四位评论人的批评，博伊德巧妙地捍卫了自己的观点。在结尾部分，他对西布莱特的批评进行了尖锐的反驳。他坚持自己那种更为乐观的看法，并以一个非常有趣的现象为例，至于这个例子是什么，在此我先卖个关子。

博伊德和里克森，以及其他合作者，比如人类学家约瑟夫·亨里奇（Joseph Henrich）和萨拉·马修（Sarah Mathew）所开创的这项研究，为重新思考人类本性、社会秩序、人类进步等一系列基本问题打下了良好的基础。这些旁征博引的论述闪烁着迷人的学术光芒，并对当前人类所面临的“制度设计与变革”等问题提出了深刻的建议。我希望，这部内容丰富的作品能够进一步激发起学者们的探索之心，并能帮助他们提出更加具体的、明确的社会改革方案。

目录

上篇　演讲篇

01　人类的独特之处（一）：模仿学习
饿死在丰裕之地
人类真正的独特之处
智慧给不了全部答案
为什么文化传承的知识如此有用
人们知道自己为什么这么做吗
盲目模仿的进化
模仿并没有那么盲目
难以理解的因果关系
文化的进化速度比基因更快
为什么人类的适应能力如此强大
过往的“文化幽灵”
人口规模与文化适应
社会学习是否会促进复杂行为的累积性进化
文化的适应性

02　人类的独特之处（二）：社会规范
专业化分工和交换活动
这一切并不是靠你个人创建起来的
小规模社会中的大规模公共产品生产
非亲缘性合作的人类
互惠利他主义并不是解决之道
制裁与互惠
图尔卡纳人为何而战
为何要实施惩罚
稳定的社会规范
文化的群体选择
为什么许多人拒绝接受文化的群体选择推论
文化的群体选择并非唯一的解释
人类的小规模合作
规范调控人类的小规模合作
第三方的监督和制裁可以稳定互惠关系
规范心理是如何得以进化的

文化让人类成了与众不同的生物

中篇　评论篇

03　模仿、哈耶克理论和文化学习的意义
超级大脑与精准模仿
哈耶克与对传统的尊敬

04　适应不需要洞察力？
个体与社会学习
对博伊德构建的模型的分析
关于规范的不同思考
以进化的角度看待人类社会生活

05　文化进化中的推论与假设检验
规范的制约
文化群体选择的假设检验

06　适应性、合作性、操控性和敌对
通过直接制裁来监督对规范的遵守情况
规范沟通和规范强化的重要性
规范强化过程中的利益冲突

下篇　回应篇

07　文化、观念和决策
回应艾伦·奥尔
回应金·斯特林
回应鲁思·梅斯
回应保罗·西布莱特

致　谢
注　释
参考文献



扫码下载“湛庐阅读”App，
搜索“人类的价值”，
获取作者演讲等精彩内容。

A DIFFERENT KIND OF ANIMAL

上篇

演讲篇

01
人类的独特之处（一）：模仿学习

饿死在丰裕之地

1860年，墨尔本市的精英人士组织了一支前往澳大利亚内陆的探险队。对于当时的澳洲白人来说，内陆还是一片未知之地。这些白人队员们的动机各不相同。有些人希望能够在澳大利亚与爪哇岛之间找到一条架设电报线的路线，从而将两地连接起来，并由此将全世界连接起来；而另一些人则是想和阿德莱德市的精英们一争高下。后者在一年前组织过一场类似的探险活动，但以失败告终。

探险队雇用了罗伯特·伯克（Robert Burke）当他们的领队，他是一名勇敢的退伍军人。此外还有18名队员，其中包括威廉·威尔斯（William Wills），一位科学家兼制图员。8月20日，探险队出发了，带着26匹骆驼、23匹马、够吃两年的食物，以及很多维多利亚时代的行李，譬如橡木炊具。包括伯克和威尔斯在内的先头部队于11月11日到达了库珀河（Cooper Creek），这是一条位于墨尔本市以北600千米的季节性河流。探险队在那里驻扎下来，等待后续队伍赶来会合。

到了12月中旬，后续部队依然未赶到，而伯克不愿再等下去了。他与威尔斯、查理·格雷（Charlie Gray）以及约翰·金（John King）再次踏上征途，希望能够到达卡奔塔利亚湾（Gulf of Carpentaria），并在3个月之内返回库珀河。伯克让队伍的其他人在库珀河驻扎到次年的3月15日。不幸的是，他们历尽千辛万苦，花了4个月才到达目的地。在归途中，格雷又遭遇了不测。当伯克、威尔斯和金于次年4月中旬返回库珀河时，营地已经被废弃。一行人精疲力竭，补给也所剩无几，很难再走上600千米返回墨尔本市。

就在此时，探险队遇到了一群扬德鲁万塔（Yandruwandha）(1)人，因处境糟糕而获得对方给予的6千克鱼。几周之后，他们又遇到了一群扬德鲁万塔人，并受到了邀请，到了对方的居住地，并在那里获得了更多的鱼，以及一种由澳洲大柄苹（nardoo）这一水生蕨类植物的种子[1]做成的糕点。这三个白人非常喜欢这种糕点，甚至觉得，现在要想活下去，就必须先学会把澳洲大柄苹的种子磨成粉。不过，他们当时并不知道这种粉是由哪种植物的种子磨制而成的，而且此后很久，一直都没有再遇到扬德鲁万塔人。

经过两周令人绝望的寻找，威尔斯等人终于找到了澳洲大柄苹，并采集了足够多的种子，磨成了粉。然而，尽管有足够的澳洲大柄苹吃，

但他们的身体还是变得越来越虚弱。到了7月初，伯克和威尔斯相继去世。最后，一群扬德鲁万塔人救下了金，给他食物，照料了他几个月，直到9月救援队伍赶到。[2]

为什么伯克和威尔斯会饿死在一片丰裕之地？至少在扬德鲁万塔人看来，那里可谓物产丰富。这个问题实际上与本文所要回答的，更为宏大的关键问题一脉相承，也就是人类何以能成为如此独特的物种。500万年前，我们的祖先也不过是一群再平常不过的猿类；而在今天，人类这种生物俨然已经成为世上其他生物的“统帅”。地球上的每一个角落几乎都有人类的影子；人类在数量上超过了陆地上大部分的脊椎动物，所拥有的资源超过了其他所有生物，社会关系也比其他所有生物都广泛。之所以会有这样的变化，是因为人类学会了以文化的方式来适应环境，并不断积累着那些对生存至关重要的信息。

在扬德鲁万塔人看来，澳大利亚内陆是丰裕之地，这是因为其文化积累了极为丰富的“如何在那里生活”的知识；而伯克和威尔斯之所以会饿死在那里，是因为他们没有接触过这方面的知识。本文旨在说明人类在进化过程中逐渐拥有了以文化的方式来适应环境的能力，而这种能力不论好坏，终究让人类成了一种独特的物种。

人类真正的独特之处

大部分进化心理学家承认，人类在很多方面都与其他物种不同：人类比其他生物更擅长使用工具；人类语言所能传递的信息比其他动物语言多得多；人类大脑相对于身体的比例，也要比大多数动物的高一些。可是，这些又有什么意义呢？很多动物都具有自身特性：大象拥有长而灵活的鼻子，蜂鸟可以在空中盘旋，蓝鹀懂得通过观察星星的位置来为飞行导航。长期以来，人们都习惯性地认为，人类拥有某些与其他动物不同的特性，因此优于其他动物。然而，很多所谓的特性已被证实绝非人类所独有。生物学家最喜欢反驳“只有人类才能做某某事情”之类的观点。制作工具、使用语言、种植作物，以及文化、教育、战争等，都曾被认为是人类的“专属”行为，后来也都遭到了驳斥。

那么，如何才能知道人类与动物之间的关键差异呢？人类真的是独一无二的吗？还是说，人类并没有我们所认为的那么独特？我想，对于这方面的研究，最好借鉴一下动物学标准，也就是在对其他物种进行比较并判断其生态重要性时所采用的标准，比如物种分布情况、物种生物

量、能量处理方式等。如果人类在这几个方面真的与其他物种有所不同，那么就说明人类在适应环境方面确实有独到之处，也说明这样的探索是值得深入的。

先来看物种分布情况这个标准。一个物种的地理分布是指它所占据的区域，而一个物种的生态幅度是指其栖息地的集合。这两个指标都很有用，因为通过它们可以大致推测出一个物种的适应能力处于何种水平。在其他条件不变的情况下，一个物种分布得越广，意味着越能在不同的环境中正常生活。在所有陆生脊椎动物中，人类的地理分布和生态幅度都是最广的。大部分陆生脊椎动物都只分布于某个大陆的某些地区，且栖息地也不会太多。人类的近亲，猿类就是最好的例证。猩猩和长臂猿只存在于东南亚的密林中，大猩猩和倭黑猩猩只生活在非洲的热带雨林中，黑猩猩只生活在非洲的热带林地中。另外，大型的捕食动物分布得最广，例如灰狼，广泛分布于欧亚大陆和北美大陆除热带以外的地区。[3]人类则生活在除南极洲之外的所有其他大陆上。

你可能会认为，人类是到了近代，借助农业和工业的发展才得以扩张到地球的每个角落的。然而，实际情况并非如此。在全新世（the Holocene）(2)之初，也就是10 000万年前，这些狩猎-采集者就已经广泛分布于除了南极洲和几个偏远岛屿之外的所有陆地区域，而且栖息地的环境千差万别，从潮湿的非洲热带雨林到极度干燥的中亚沙漠，再到北冰洋的冰岸，不一而足。

我有时候也会拿啮齿目动物和人类做比较：啮齿目动物不也是遍布全球吗？是的，它们确实遍布全球，不过却包含了2 200多个物种，而其中任何一个物种的分布范围都没那么大。人类则不同，现代智人这一物种几乎分布于地球表面的每一个地方。褐鼠看上去虽是一个例外，但事实并非如此。它们的分布范围几乎和人类一样广泛，遍布其他各个大陆。然而，事实上，直到中世纪，褐鼠都只分布于中亚，后来“搭乘”人类的交通工具，才逐渐泛滥于世界各地。[4]褐鼠和螨虫、肠道寄生虫等人类共生物一样，则都生活在人居环境之中。

如果用另外两个动物学标准来衡量，得到的结果也是一样。在所有的脊椎动物中，人类的物种生物量[5]仅少于我们驯养的家禽，同时也是所有野生陆生脊椎动物的物种生物量之和的数倍。[6]人类如此多的物种生物量，不仅仅是农业生产和工业发展的结果。据估计，在全新世之初，地球所承载的狩猎-采集者数量约为7 000万。[7]这么庞大的数值是值得关注的，因为在很多环境里，这些狩猎-采集者都是顶级的，能够猎

捕领地范围内体形最大的动物。当然，顶级狩猎者通常在数量上要比猎物少很多。据估计，在全新世之初，狮子作为顶级狩猎者之一，纵然分布范围较广，但数量也只有700万只左右。

这种生态学意义上的成功，源自环境适应性。啮齿目动物几乎可见于人类的每一处栖息地，不过这类动物有许多亚类，各种亚类的适应能力不尽相同，虽然都能适应某种特定的环境。[8]河狸拥有带蹼的脚和扁平的尾巴，更适应水中的生活；美洲飞鼠拥有巨大的飞膜，能够从一棵树上滑翔到另一棵树上；更格卢鼠（kangaroo rat）拥有独特的肾脏，就算不喝水也能在沙漠中生存；盐沼小鼠则可以喝盐水。然而，它们的种种差异并不仅仅停留在生态学层面上。

不同的物种会采用不同的社会行为来适应独特的环境。有些动物遵循一雌一雄制，而有些则很混乱；有些物种习惯独来独往，而有些物种，如河狸、草原土拨鼠、裸鼹鼠（naked mole rat）等，则以集体合作的方式生活。与所有这些物种都不同，人类是通才，不仅能够适应各种环境，还能总结衍生出各种地方性知识和专业化工具，以及各种社会安排。

最能表现人类适应能力的行为就是加工各种东西，但这种加工能力并不能让我们适应各种不同的环境。很多动物都能加工东西，比如鸟能筑巢，河狸能筑河坝，白蚁能打造蚁丘等，而且加工能力在这些动物的生活中扮演着重要的角色。这些动物制造品，就技术复杂性而言，完全可以与几千年前的人造物品媲美。织巢鸟的悬巢非常漂亮，而且复杂程度一点也不亚于许多狩猎-采集者修筑的茅草屋。然而，每一种织巢鸟只会搭建有限的几种鸟巢。[9]人类真正的独特之处在于，我们能够制造出各式各样的东西，从而完美地适应独特的生存环境。人类会根据地域环境的不同而建造不同的住所。我们的房子可能是用冰雪筑成的，也可能是用泥土、石头、茅草和木头搭建而成的。不仅如此，人类还会做出各种各样的设计，以便更好地适应当地环境。

总而言之，人类在适应不同的环境方面，表现出了超乎寻常的能力。那么，接下来的问题是，人类为什么会拥有这种能力？

智慧给不了全部答案

一个颇为自得的回答是：人类之所以如此成功，是因为人类更聪明。大部分研究人类进化的学者都认同这一点。另外，更为复杂的工具

和狩猎-采集方式，以及符号行为都可以证明，人类的认知能力是在不断提升的。[10]考古学家林恩·沃德利（Lynn Wadley）及其同事发现，7 000年前，居住在非洲南部海岸斯布都洞穴（Sibudu Cave）的人们，在制作箭矢或长矛时，会使用黏合剂将锋利的石片固定在木柄上。[11]实验发现，这些黏合剂是由金合欢树胶和代赭石混合加热而成的，而且比用其他方法制作出来的黏合剂更好用。

他们总结道：

> 我们的实验表明，这种黏合剂的体积吸收总系数（KYA）达到了70……这些制作者的技术相当娴熟，他们不仅对这种黏合剂材料的特性了如指掌，而且还会有意识地操控这些特性。[12]

这段话包含了两个观点。第一个观点是，这些制作者的技术非常娴熟。之所以这么说，是因为在那个时候，制作和使用黏合剂是一件非常复杂的事情。第二个观点是，他们的技术之所以会如此娴熟，是因为他们非常聪明，能够理解这种黏合剂的特性。但是，纵然第一个观点是对的，第二个观点却未必正确。

大部分研究人类进化的学者对智慧转化为行动的具体过程都不甚了解，也不清楚如何用进化来解释智慧。最简单的观点是，人类在适应环境方面与其他物种并无二致，无非是适应得好一些罢了。其他物种在适应环境时所采用的方式主要有两种。一种方式是通过基因变异来适应环境，其典型实例是，居住在高纬度的动物在体型上要比居住在低纬度的同类更大一些。这种差异就是适应性的结果，因为大体型更容易保存热量。自然选择让生活在不同纬度的相同物种出现了基因差异，而这种差异影响了物种个体的体型。与此类似，自然选择也会影响行为。例如，某种环境中的适应性行为是投掷长矛，而另一种环境中的适应性行为则是伏击猎物。但是，自然选择是一个相对缓慢的过程，因此基因差异并不能解释为什么居住在斯布都洞穴里的人会用金合欢树胶制作黏合剂，而澳洲土著居民则是用三齿稃树脂来制作黏合剂。

这就涉及了另一种适应方式：动物个体在生命过程中通过行为和形态上的变化来适应当地环境。对脊椎动物而言，学习尤为重要。动物个体不仅要学习去哪里寻找食物，以及寻找什么样的食物，还要学会分析谁会允许自己与之共享食物来源地。总之，要学会灵活调整自身行为以适应环境。因此，它们总能迅速地适应不断变化的环境，并能在多种类

型的栖息地中生存。

非常重要的一点是，对于其他物种来说，学习通常只存在于个体层面。个体会通过自己的方式来学习所需要掌握的知识和技能，而不会向同类学习。的确，很多脊椎动物都拥有一些简单的传统行为，而且这些传统行为会通过模仿等社会学习机制得到强化。然而，不管在何种情况下，这些传统行为都包含了个体采用自己的方式所掌握的行为。[13]社会信息非常有用，但并非必不可少。毫无疑问，就算不借助任何社会信息，个体也能通过学习来做出高度适应性的行为，比如制作加工非常复杂的用品。织巢鸟在没有见过同类如何筑巢的情况下也能建造出非常漂亮的巢穴。[14]刚刚长大的织巢鸟在筑巢方面做得并不是很好，但随着经验的积累，其筑巢的“手艺”会得到改进。人工饲养的䴕形树雀会学习如何利用工具把虫子从树木的内皮中掘取出来；而对于有经验的树雀来说，把树皮扒开是完全没有必要的。[15]人类的学习过程看起来也大抵如此。

接下来我会尽力向大家证明，这么说是错误的。人类并非像其他物种个体那样，单纯依靠自身能力来适应环境。超常的智慧使人类能够适应更多类型的环境，但还不足以让人类聪明到能够以个体的方式来解决现代人类所面临的所有适应性问题，毕竟人类已遍布全球。任何狩猎-采集群体所使用的工具、寻找食物的技术、生态知识，以及社会安排都非常复杂，绝非个体所能创造。人类之所以能够掌握在不同环境中所需要掌握的所有知识和技能，是因为懂得从他人那里获取信息。在向同类学习这个方面，人类比其他任何物种做得都要好。另外，同样重要的一点是，出于某种心理，人们很愿意向他人学习，甚至在不了解对方初衷的情况下也是如此。这种心理机制使人类积累起了极为丰富的适应性信息，其“实力”远远超过了个体的创造能力。对于人类的适应能力来说，累积性的文化进化是非常关键的驱动力。人类如果未能通过文化的方式来适应环境，那就不会成为一个独特的物种。

这让我们再次联想到命运多舛的伯克-威尔斯探险队。在我看来，这个故事不过是所谓的“欧洲探险者迷失经历”的例证之一罢了。在过去的几个世纪里，类似的故事一再上演，结局也都大同小异。[16]一小队欧洲探险者在陌生地带身陷困境，而当地土著却过得顺风顺水。尽管付出了巨大的努力，也花费了大量的时间去学习，但这些迷失的探险者仍然无法找到可以用来果腹的食物，到最后往往只有死路一条；有幸活下来的人，通常也是有赖于当地土著的热情帮助。

伯克-威尔斯探险队的队员之所以会丧生，并不是因为不够聪明，[17]而是他们从未接触过扬德鲁万塔文化中的那些传统知识。正是这些知识让扬德鲁万塔人能够在库珀河流域生存下来。不妨从大柄苹开始讲起吧。这些白人没能依靠自己的探索发现大柄苹，但在看见扬德鲁万塔人将大柄苹的种子磨成粉之后，也学会了做同样的事情。然而，他们最终还是死了。

威廉·威尔斯当时觉察出事情不太对劲，曾在日记中写道：

> 虽然胃口很好，大柄苹也非常好吃，但它似乎无法给我们提供任何营养。我比任何时候都虚弱。这里的鸟儿如此怕人，根本捉不到。即便有不少鱼，但我仍然不太清楚，在只有大柄苹的情况下，我们还能搞出什么花样来。能够拯救我们的，除了天大的好运，别无其他。至于我自己，大概只能再活四五天吧。[18]

然而，威尔斯并不知道，其实大柄苹含有一种酶，这种酶会降低肠胃对维生素B_1的吸收。这也就是说，虽然这些食物能提供足够的卡路里，但他们仍会饱受脚气之苦，而且很可能得了坏血病。[19]在制作大柄苹时，扬德鲁万塔人会用大量的水反复冲洗这种植物。很多植物都含有毒素，而以植物为主食的生物自然拥有各种处理毒素的方法。对于这些欧洲人而言，就算知道大柄苹有毒，要学会去除毒素也并非易事，毕竟从未接触过扬德鲁万塔文化中的这类传统知识。如此一来，大柄苹的坏处远远超过其好处，因为它有毒。

即便如此，在那个环境中，还有其他一些食物可以吃。扬德鲁万塔人捕了很多鱼，还给了这几个欧洲人6千克鱼。罗伯特·伯克和威尔斯既然知道大柄苹不管用，那为何不去捕鱼？我们不太清楚，但或许存在这样一种可能：许多澳洲土著会用渔网来捕鱼。居住在库珀河流域，扬德鲁万塔部落北面不远处的另一个土著部落迪亚里（Diyari）就是这样捕鱼的。[20]这也就是说，伯克和威尔斯应该看到了扬德鲁万塔人使用渔网捕鱼，那么他们为什么不做一个渔网呢？这又是另外一个难题了。要想制作渔网，他们需要先学会利用当地的材料制作绳索。在澳大利亚的这一地区，土著会利用一种被称为“verbine”的，开紫色小花的灌木的树皮和根部来制作绳索。[21]在收集好这种植物后，再将纤维物质从中分离出来。为此，整株植物会被泡在水里好几天。然后，树皮会被撕成条状，经过反复敲打和揉搓而变得十分柔软。最后，晾干即可编成绳索。土著

会将这些纤维放在大腿上揉搓，然后拧成双股线。拧好线之后，还需要知道如何将其织成渔网。居住在库珀河流域的人们拥有两种织网技术，分别针对不同类型的捕鱼活动。

对于扬德鲁万塔人来说，库珀河流域是一个丰裕之地，因为他们的文化保留了大量的传统知识，这些知识可以帮助他们在此生活下去。一部关于扬德鲁万塔人的“自然历史书”可能会有上百页，并分为“如何游戏”“有效的打猎技术”“如何找到水源”“如何处理有毒的蕨类、薯类和苏铁类植物”等板块。在考古学家那里，澳洲土著以技术简陋而“著称”。然而，这样一本“技术操作指南”可能会包含如何制作和使用渔网、篮子、房屋、回力标、灭火器、梭镖、毒药、黏合剂、盾牌、树皮舟、石制工具等。此外，我们将会在下文中看到：合作，在人类的生存中扮演着非常关键的角色。要想成为一名合格的扬德鲁万塔人，还需要掌握其“社会政策及相关程序”“语法和词典”，以及“信仰、故事和歌曲”等，这些内容若是打印出来，大概也会有厚厚的一摞纸。

为什么文化传承的知识如此有用

要知道，人类与其他动物在适应环境时所采用的方式是不同的。我们不用重新探索和学习所需的知识，因为文化传承能带来很多适应性知识。那么，这些传统知识是如何发挥作用的？人们如何确信传统知识是有用的呢？有关这方面的研究少得可怜。在我看来，人们通常会这样想：创新是很难的。发现大柄苹有毒并不是一件容易的事，而找到去除毒素的方法更是难上加难。好在，确定一种行为是否有益相对比较容易。因此，一种发明一旦出现，只要被确认是有益的，便会被传播开来。回想一下考古学家林恩·沃德利的观点，那些在7　000年前就学会了制作黏合剂的非洲土著人拥有非常娴熟的制作技术，不仅完全清楚那些黏合剂材料的特性，还会有意识地操控这些特性。[22]

因此可以说，文化就像一座图书馆。这间图书馆里保存的都是以往创新所得的有用知识。图书管理员需要判断并决定哪本书该放在书架上，但书的内容则取决于作者的能力，而非图书管理员的能力。技巧娴熟的作者通常能够写出更好的书籍。同样的道理，虽然文化会保留各种创新，但一种发明的有用性实则取决于创新者的头脑。于人类而言，没有卓越的认知能力，就无法保存任何复杂的适应性创新。

对人类进化学感兴趣的人，对文化的看法都大抵如此。在一篇经典

的文献中，美国人类学家约翰·图比（John Tooby）和伊文·德沃尔（Irven DeVore）皆认为，人类已经步入了认知生态位（cognitive niche）。这是一种解决问题的方式，其他动物无法企及。[23]人类能够在陌生环境中发明新工具并找到合适的技术，因为人类比其他动物更善于因果推理，以及其他各种形式的心智模拟。懂得相互学习，意味着每个人都不需要单独发明自己所需要的各种事物，因而创新的成本便得以分摊。[24]

对于支持进化论的研究者来说，这种方法极具吸引力，因为在他们看来，解释清楚自然界中的设计至关重要。[25]这里的“设计”指的是复杂的、不可解的事物构造；这种构造能够实现某些功能。例如，眼睛这种器官明显就是为了看东西而“设计”出来的。眼睛要看到东西就必须保证水晶体、虹膜和视网膜等按照一定的结构组成，而这种结构绝不可能是偶然出现的。设计还适用于一些复杂的行为，比如处理大柄苹和制作渔网。自然界中之所以存在设计，唯一的原因就是自然选择，就连人类大脑都是自然选择的结果。由此可见，文化是各阶段创新的集合，而这些创新之所以能被保存下来，是因为人们欣赏其价值。

在人类学、社会学和心理学中，还有一种颇具影响力的看法：人们信奉不同的信仰，并践行相应的价值观。这些信仰和实践可能是合乎情理的，也可能不是。而且，或许人们并不清楚，为什么一种信仰和实践比另一种信仰和实践更好。人们会相信自己愿意相信的事物，按照自己的一贯风格行事，因为同一文化环境中的其他人都是这么做的。

那么，究竟哪种观点才是正确的？个体采纳某些观念和行为，是因为深知其益处，还是因为看到周围的人皆是如此？如果后者是正确的，那么凭什么说某些观念和实践就是有用的？

人们知道自己为什么这么做吗

扬德鲁万塔人知道自己为什么要这样处理大柄苹吗？我们永远也不会知道答案了，因为扬德鲁万塔人及其生活方式在罗伯特·伯克和威廉·威尔斯死后的几十年中渐渐消失了，枪炮和病菌夺走了他们的性命。不过，我们对当代一些种群的食品处理技术的文化进化做了些了解。

斐济群岛上的食物禁忌就是一个很好的例子。人类学家约瑟夫·亨里奇及其同事，包括我在内，曾对斐济亚萨瓦岛（Yasawa Island）上几个村庄的生活方式进行了长期研究。那里的人们过着自给自足的生活：

80%的卡路里来自白薯、木薯、香蕉，以及其他一些田间作物，还有鱼、海龟和一些海洋无脊椎动物；20%的卡路里来自购买的食物，主要是糖、面粉和食用油。在斐济人的食物中，大部分海产品都含有毒素。

亨里奇所说的是一种被称为雪卡（ciguatera）的毒素，这是一种由生长在死礁上的鞭毛藻所产生的化学物质。[26]雪卡毒素溶于油脂，会积存在食用鞭毛藻的体内。食物链顶端的动物，如梭子鱼、鲨鱼、海鳝，以及一些寿命较长的生物，如海龟等，体内累积的雪卡毒素最多。雪卡毒素会对人体造成严重的、持续性的伤害，中毒症状包括关节疼痛、忽冷忽热、腹泻和呕吐。这是困扰整个大洋洲的健康问题之一。医学研究表明，雪卡毒素对胎儿和吃母乳的婴儿危害最大。我们在走访亚萨瓦岛居民时发现，大部分居民都认为孕妇和哺乳期妇女应该避免食用那些容易感染雪卡毒素的海产品，比如鲨鱼和梭子鱼一类的大型海洋捕食者，以及海龟一类的长寿生物，详见图1-1。这种观念还被赋予了道德色彩，事实上，斐济人不允许孕妇和哺乳期妇女吃这些食物。

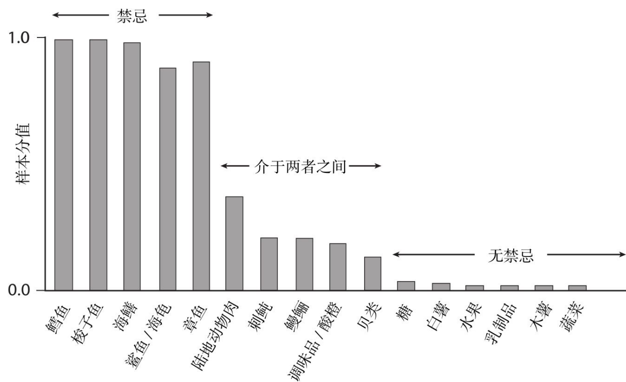


图1-1　对**70**名斐济妇女的食物禁忌调查。这些妇女都认为，孕妇不应该吃表中所列的禁忌食物。研究者列出了一长串食物的名称，然后问这些妇女“这些食物是否适合孕妇食用”。在询问当地人“哪些食物是孕妇应该避免的”时，也得到了类似的结果。

From Henrich & Henrich, 2010.

因此，我把这种信仰称为禁忌。斐济人将此称为“tabu”，发音与英文中的“tamboo”（禁忌）一样，而且与库克船长（Captain Cook）所引入的词汇汤加（torgan）也很接近。多项证据表明，这些禁忌是环境适应性方面的传统文化知识，它降低了该人类种群中部分弱小群体感染雪卡毒素的风险。

村民们不仅知道吃鱼可能会生病，还能识别出这种特殊疾病的症状，并称之为“ika gaga”，字面意思是“鱼毒”。他们所描述的症状与西

方医学界确诊的雪卡毒素的中毒症状几乎一样。村子里的大部分人在一生中至少会中一次雪卡毒素。他们知道，鱼的体型越大，中毒的风险越高，[27]但是并不相信食用海龟或海鳝会导致雪卡中毒。因此，这些禁忌并非源自“吃鱼中毒”这种一般信仰。事实上，它所针对的只是孕妇和哺乳期妇女。这些禁忌会对孕妇和哺乳期妇女的生活产生重大的影响，因为这些“不能吃”的食物实际上是斐济人日常饮食中的重要食材，而且比其他鱼类和海洋无脊椎动物的油脂含量更高，是斐济人摄取油脂的重要来源之一。当然，它们深受好评还因为味道鲜美。关于这一点我可以证明，在吃了数周的木薯和珊瑚鱼之后，你一定会对海龟的美味赞不绝口。这些禁忌所限制的是那些极易感染雪卡毒素的物种，而就其他一些物种而言，比如鹦鹉鱼和独角兽鱼，则更常被食用，不存在什么禁忌。

大部分人都是从其他人那里得知这些禁忌的。亨里奇和同事们询问了一些斐济妇女“是如何知道这些禁忌的”，结果见图1-2。有5%的人说，她们知道“孕妇和哺乳期妇女需要忌口”是与自己的切身经验有关；95%的人表示是从其他人那听说的。这部分女性口中的“其他人”大部分是指家庭成员，但也有大约25%的人是指那些“有智慧的女性”，比如传统医疗、孕产以及女性养护等重要领域内的专家。

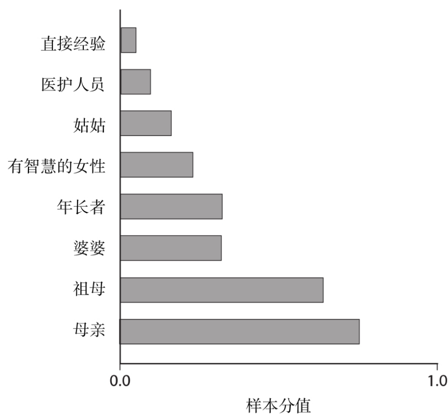


图1-2 对**44**名斐济人的调查结果。他们所认同的“某些食物不适合孕妇和哺乳期妇女食用”的信仰来自何处？

From Henrich & Henrich, 2010.

那么，为什么斐济人会相信父母、亲戚，以及有智慧的女性所告之的这些经验呢？依照“文化是图书馆”这个观点，答案如下：这些学习对象会给出极好的理由让他们相信孕妇和哺乳期妇女不应该吃海鳝、海龟等食物，并遵照这些信仰行事。这些信仰需要具备足够的说服力，只有

这样才能让一个品尝过美味海龟的妇女，在其怀孕时顾及这些食物给未出生的孩子所带来的不利影响，从而选择天天吃煮熟的木薯。这些信仰必须具有强大的说服力，只有这样人们才敢指责一个吃海龟的孕妇，即便这种指责可能会冒犯她和她的亲属。食物和疾病之间的因果关系将会得到强化，因为事实证明这种关系是真实存在的。例如，村子里的某个人注意到，某位妇女在不知道自己已怀孕的情况下吃了不少海龟，后来生出的孩子在发育方面确实存在一些问题。

但是这种观点并不符合实际情况。亨里奇及其团队调查了很多妇女，询问她们“一个孕妇或者哺乳期妇女如果吃了一些禁忌食物，会发生什么事情”，答案可谓五花八门。一些妇女不知道，另外一些则觉得什么也不会发生。大部分人都认为，那样做会对婴儿健康产生不利影响，但不利影响到底是什么，又众说纷纭，“皮肤粗糙”和“体味难闻”是其中比较有代表性的两个答案。这种模式与传统观点是一致的，也就是说，人们之所以会信奉某种禁忌，只是因为大部分人都相信它；如果要他们做出解释，得到的答案通常只是，不遵守这种禁忌便会遭遇一些不好的结果。值得注意的是，那些孕妇和哺乳期妇女就算无法对禁忌做出合理解释，也依然会放弃那些美味的、令人垂涎的食物。

这里还有一个例子。我曾经对斐济这几个村庄的传统房屋进行过研究。这些房屋被称为“bure”，发音接近“booray”，功能性很强。不仅能遮风挡雨，还能为人们提供烹饪和社交活动场所，以及保护家庭隐私。这些房屋的形状和大小各不相同，但结构都大致相同（图1-3）。当然，这也是斐济群岛上最流行的建筑样式。

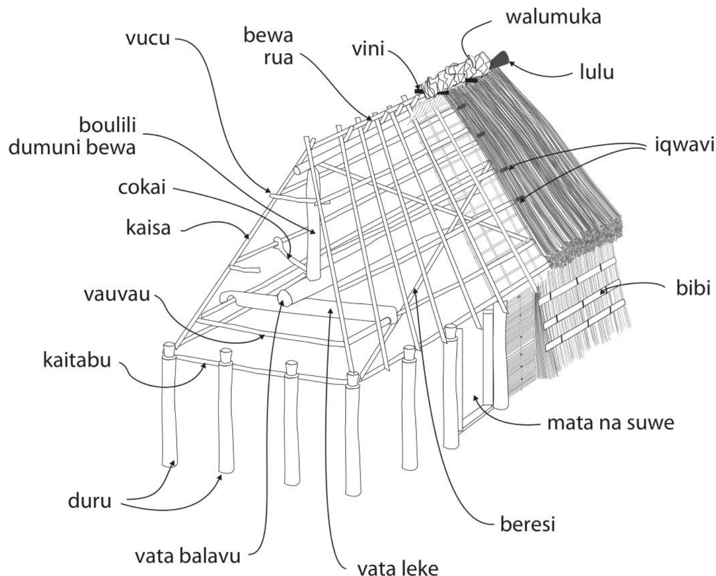


图**1-3**　斐济传统房屋的剖面图。图中的标签是亚萨瓦岛上的斐济人用当地方言写下的建筑结构名称。按比例换算后可知，标识为“**duru**”的立柱高约**1.5**米，被放置在深约**1.5**米的孔洞中。据我们所知，“**duru**”的发音为“**nduru**”，而孔洞未被明确标识出来。图为草屋的基本构架，屋内厚梁（**vata leke**）高于头顶，屋外覆盖着厚厚的茅草。（图中语言为斐济土著方言）

斐济房屋为何要按照这种模式来建造？为了找到答案，我重点关注了房屋结构，想弄清楚这种设计能否抵御每5年左右就会来一次的横扫全岛的台风。风暴常常会摧毁房屋。2009年，一场相对温和的台风过境，村子里约有15%的房屋都坍塌了；上了些年纪的人们至今还记得20世纪50年代早期的那一场台风，当时所有的房屋都倒塌了。建造这样一个房屋需要6个劳动力花上1个月的时间，而这与重新打理被风暴摧毁的田地所需要的人力和时间几乎相同。我寻访了8位村子里公认的最懂房屋建造的男性，请教了诸多房屋结构方面的问题。除此之外，我还想知道，他们是否能够解释清楚，为什么这种标准的房屋建造模式比其他的要好。

在走访快结束时，我展示了斐济群岛其他岛屿上的房屋图片，并询问，如果风暴来袭的话，这些房屋是否能够抗得住。他们都认为，其他类型的房子不太好建造，但至于为什么不好建造，却说不出理由。我想，可能是因为这些房屋与当地的房屋差别太大了。于是，我从这种草屋的原型出发，只改变其中一个特性，比如采用不同类型的木材或不同结构，然后再次询问，与标准的草屋相比，“我的房屋”有什么不同。对于某些特性，他们能够给出明确的因果解释，比如，“noko　noko”适合在旱地里做柱子，而“vau”适合在靠近海滩的比较潮湿的土地上做柱

子，但是说不清楚为什么我改过的一些结构不如标准结构好。

由此看来，亚萨瓦岛居民并不能为他们的食物禁忌提供很好的解释，且对房屋结构的认知也是不全面的。吃着自己应该吃的，按照自己的方式建造着房屋，或多或少是因为其他人都这么做。无论身在何处，人们总会受到周围人的影响。更正式地说，人们似乎本能地倾向于采纳他人的观点。那么，自然选择会青睐这种心理机制吗？

盲目模仿的进化

理论上，自然选择会青睐这种心理机制。这种心理机制会使大多数人将周围人的信仰当作自己的信仰。[28]我之所以强调“理论上”，是因为这个结论来自一套数学模型，而像我这样的研究者通常会利用它来研究进化过程。我们不妨先口头论证一下这个有趣过程的基本特性。在这里，先假设，最优行为会存在一些变化。比如，在某些环境中，孕妇不应该吃大鱼；而在其他环境中，吃大鱼可能会比较好。在某些环境中，人们注意到，那些在孕期吃了很多大鱼的妇女生出来的孩子有问题；于是，妇女们利用这一线索推断，怀孕时不应该吃大鱼。但是，还存在一种可能性：想要确定吃大鱼所带来的影响是很困难的，因为样本太小，而影响婴儿生长发育的因素又有很多。

看起来很有可能是这样：如果环境中所出现的线索是可信的，那么就使用这些线索进行推断；如果这些线索是不可信的，那么就依靠模仿。这种策略可能会受到自然选择的青睐，这样一来，个别情况下对可靠线索的使用，就会通过模仿而被传播开来，从而使最容易被观察到的行为成为最优行为。为了验证这一口头论证，需要建立一个模型，用数学形式来表示其基本假设；接着，要用数学工具来求证自然选择是否会青睐这一特性。如果最终结论与口头论证的结论一致，就说明这个论证很可能是有道理的。严谨一点的话，需要再建立一个不同的数学模型来表示这个基本假设。如果得到了相同的答案，那么就可以确定，口头论证是对的；如果得到的答案不一样，就需要找出哪里出了问题。

接下来将建立一个非常简单的模型，以证明口头论证是正确的。[29]本书的两位评论人建议我删除这一部分，因为数学会让许多人望而生畏。最终，我还是决定保留，因为让读者了解这种理论推导过程是十分重要的。大家尽力理解吧，如果觉得像是在读天书，可以直接跳到结论部分。

假设，有一大群人生活在同一环境中，而这个环境只有两种状态。为了简单起见，我把这两种环境分别称为“大鱼有益”和“大鱼有毒”。需要注意的是，这并非是指斐济的鱼类禁忌模式。[30]目的是检验口头论证的逻辑合理性。假设这些人生活在一个不断变化的环境中，若非如此，他们就没必要学习了。假设环境会在这两种状态之间不断地随机切换，在一段时间里，环境处在一种状态，而后又切换到另一种状态，且两种状态存在的时长比较平均；存在两种可能的行为——“吃大鱼”和“忌吃大鱼”。当环境处在“大鱼有益”状态时，“吃大鱼”就会给人们带来更大的好处；而当环境处在“大鱼有毒”状态时，“忌吃大鱼”就会带来更大的好处。由此带来的适应性问题就是，这些人需要确定自身处在哪种环境状态中，以便选择最优行为。

有两种信息可以帮助个体确定自身所处的环境状态。第一种信息是环境线索，个体可以通过环境线索来推断环境状态。可以将这种环境线索视为一种相关性，比如在学习者所知道的某个样本群中，一位妈妈吃的大鱼数量和她宝宝的健康状况之间的相关程度。因为是口头论证，所以这些线索难免会存在缺陷，而这些缺陷有时候会导致推断错误；不过，学习者可以通过其所含的相关信息来评定其质量。

正如图1-4所示，当处在“大鱼有益”状态时，通常为正相关；当处在“大鱼有毒”状态时，通常为负相关。因此，婴儿的健康状况和母亲所食用的大鱼数量之间的关系，如果呈正相关，也就是支持“大鱼有益”；如果呈负相关，则支持“大鱼有毒”。然而，现实世界充满了各种杂音，由此带来的一个结果是：当“大鱼有毒”时，有可能会观察到正相关关系；当“大鱼有益”时，有可能会观察到负相关关系。个体如果采纳了这些带有误导性的线索，就会做出错误的推断，以及不恰当的行为。

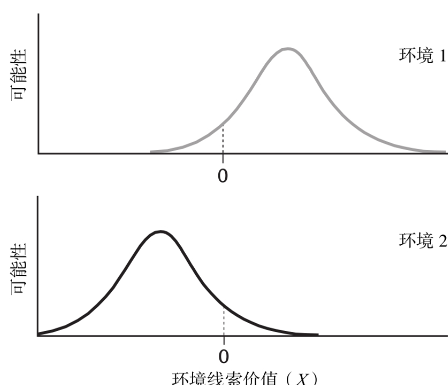


图1-4　不同的线索价值x在“大鱼有毒”和“大鱼有益”状态时的可能性。可以把这些线索价值看

作所能观察到的母亲饮食和孩子健康之间的相关性。当“大鱼有益”时，正相关的可能性更高；当“大鱼有毒”时，负相关的可能性更高。这意味着，当观察结果为正值时，环境处于“大鱼有益”状态；结果为负值时，环境处于“大鱼有毒”状态。然而，这种推断有时候也会出错。不过，线索的绝对值越大，在依靠该线索进行推断时，出错的可能性就越低。

另外，大家要注意，图1-4中的函数分布是偏移的。上图表示，当环境处在“大鱼有益”状态时，个体不太可能会观察到一位母亲食用的大鱼数量与其后代健康状况之间的负相关关系。同样，下图表示，当环境处在“大鱼有毒”状态时，个体不太可能会观察到正相关关系。这意味着，当相关性足够显著，比如呈绝对值时，是不太可能推断失误的。

第二种信息是个体观察到的群体中一部分人的行为。这些人常被称为“榜样”。假设，以平均值为标准，好的行为更常见。因此，如果大部分人食用大鱼，那么这个线索即表明“大鱼是有益的”；如果大部分人不吃大鱼，那么则表明大鱼是有毒的。但是这种线索也有可能出错。比如，很有可能会出现这样的情况：一个群体中的大部分人都认为大鱼是有毒的，但一个学习者偶然遇到了另一群人，其中大部分人都认为大鱼是有益的。还可能会发生这样一种情况：环境已经改变了，大鱼现在变得有毒，但大部分人依然觉得大鱼无毒。

因此，个体不得不在无法准确推断出最优行为的情况下决定该如何行动，必须赌一把。贝叶斯决策理论（Bayesian Decision Theory）是代数的一个分支，主要用来进行风险决策，让自身的平均收益最大化。假设一个单独的环境线索价值x，表示母亲的饮食和其后代健康状况之间的相关性。个体观察到，在一个由n个个体组成的群体中，有j个榜样都在吃大鱼。因此，运用这个理论，如果$j-n/2>-gx$，那么选择吃大鱼就可以使预期收益最大化。在这个公式中，g是一个参数，表示个体对两种信息的权衡。虽然g值非常重要，但暂且把它放在一边，先来关注这个不等式的其他部分。

当大部分榜样都吃大鱼，即吃大鱼是最优行为时，这个不等式的左边将会是正值。因为g值总被认为是正值，因此，当吃大鱼和婴儿健康之间为正相关，即x值为正时，不等式的右边将会是负值。两种信息都指向同一方向，个体吃大鱼可以让个体预期收益最大化。如果大部分榜样都不吃大鱼，即$j-n/2<0$，吃大鱼和婴儿健康之间是负相关，即$x<0$，那么不等式会告诉你，以平均值为标准的话，吃大鱼将会降低收益。

当两种信息产生冲突时，会发生一些有趣的事。例如，当10个榜样中有7个吃大鱼，即$n=10$，$j=7$时，不等式的左边$j-n/2$等于2。大部分

人都吃大鱼，这样的观察结果表示大鱼是无毒的。然而，在吃大鱼和婴儿健康之间又存在着负相关，相关值为-0.5，这意味着相反的推断。在这个数字样本中，如果$2>0.5g$，那么不等式表示的是大鱼无害。因此，你的决策取决于参数g，它表示个体对环境线索和社会线索的权衡。g值较小，意味着个体认为社会线索更重要；g值较大，意味着个体认为环境线索更重要。在这个例子中，如果g值小于4，那么个体会忽略环境线索，模仿大部分榜样的行为；如果g值大于4，那么个体会忽略周围人的行为，依靠环境线索行事。

在这个简单的模型中，参数g代表着个体如何通过学习心理机制在社会线索和环境线索之间做出选择。极小的g值意味着个体极度依赖模仿他人。为了确定自然选择是否会导致个体的“盲目模仿”，我们需要让参数g不断进化，看看到底是什么条件导致了g值变小。

因此，需要建立一个模型来研究影响学习心理机制的基因与行为的共同进化。假设g值是一种基因特性，每个个体都会从父母那里遗传到一个g值；可能会因为变异而发生一些改变，但g值与性别无关。个体通过学习获得环境线索，并观察到n个榜样的行为，然后按照上述规则选择其一。这个选择会决定个体的行为，从而影响个体的收益。如果所携带的g值能够增加个体正确识别环境线索的概率，那么这部分人将会拥有更多的后代；自然选择将会使人群中拥有这种心理机制的人越来越多。然而，最佳学习心理机制同时还取决于上一代人的行为。如果大部分榜样都拥有正确的行为，更依赖社会线索，即g值较小，这种行为将会受到自然选择的青睐；如果大部分榜样的行为都不正确，那么自然选择将会倾向于不太依赖社会线索，也就是g值较大时的行为。基因影响着人们的决策，决策又影响着下一代榜样的行为。因此，基因进化和文化进化是密不可分的，这一过程被称为“基因-文化的共同进化”。人类不断重复着这一过程，直到g值的平均值不再发生变化。[31]这个数值告诉我们，自然选择是如何在漫长的历史进程中造就人类的学习心理机制的。

图1-5展示了两组不同参数被代入模型后所得到的结果。由于g的原始数值解释起来比较麻烦，因此我假设g值为一个长期平均值，然后测算在当前环境中掌握的最优行为与榜样数量之间的关系，从而找出这种可能性：当最优行为比较容易学习时，社会线索在人们的决策过程中常常被忽略；当最优行为比较难学时，大部分个体会选择模仿榜样的行为。

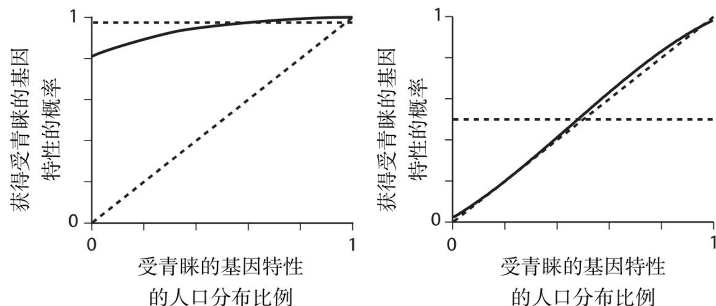


图**1-5**　当环境线索质量比较高（左图）和质量比较低（右图），人类处于一种稳定的进化状态时，受自然选择青睐的基因特性的概率。其中，环境线索质量高低是指某种比较受青睐的基因特性在人群中的分布比例。在这里，个体会关注另外三个个体的行为；其他个体数量更多的话，个体则会依赖社会线索。受自然选择青睐的某种基因特性的分布比例指的是人群中采取最优行为，也就是"当大鱼无害时食用大鱼，当大鱼有毒时不吃大鱼"的人数比例。水平虚线表示的是只利用环境线索，也就是说，只利用个体学习所能获得的受自然选择青睐的基因特性的可能性。在左图中，个体通过个体学习来选择最优行为的概率为**98%**；在右图中，概率只有**50%**。**45**°斜虚线表示的是盲目模仿上一辈人以掌握最优行为的可能性。黑线表示的是在稳定的进化状态中，使用最优的*g*值以获得受自然选择青睐的基因特性的可能性。如果个体学习是准确的，那么社会学习对个体的收益几乎没有什么影响；如果个体学习不准确，那么大部分个体都会模仿榜样所表现出来的行为。在这两种情况下，平均每经历**100**代人，环境状态就会转换一次。

Redrawn from Perreault et al., 2012.

来看看这意味着什么吧。个体拥有一种学习心理机制，这种机制告诉他们最优行为取决于在所处环境中观察到了什么。当个体学习比较困难时，环境线索能够让50%以上的人做出正确的行为，让近一半的人做出错误的行为。几乎所有人都忽视了这一论断：自身经验才是最优的，他们只是在模仿周围人的行为。换句话说，在这个模型中，个体在本能上倾向于模仿他人行为，当然也会遵循自身经验所带来的启示，但这种动机必须足够强大，他们才能克制住模仿他人的倾向。如果像通常所假设的那样，与模仿相比，个体学习的成本更高，那么自然选择就会更青睐模仿。

运用这种决策规则的群体所获得的平均收益会比完全依赖个体学习的群体获得的更多，而且在个体学习比较困难的情况下，这种差距还会进一步扩大。在这个模型中有两种错误来源。首先，个体在学习时如果依据的是相关环境状态中的部分信息，那么就会很容易出错。社会学习使个体拥有一定的选择权，要么依赖高质量的环境线索，要么模仿他人；采取模仿行为的人越多，个体学习所创造的信息质量就越高。其次，社会学习存在的问题是，如果环境发生变化，那么个体所模仿的某个行为可能会不再适应环境。这意味着，个体的模仿行为越多，群体在

追踪环境变化方面就越不准确。因此，在个体学习非常困难，环境变化比较缓慢时，模仿将会带来最大的收益。

这一结果表明：在人类适应环境的过程中，文化扮演着非常关键的角色。大家可以把“平均收益”看作“成功适应”的指标之一。这个模型说明了，当人们难以发现有益的创新时，文化在促进人类的环境适应性方面作用显著，然而人们一旦发现了有益的创新，这些创新便会长期有用。在我看来，在缓慢变化的环境中不断解决难题，正是人类进化的本质。个体很难弄明白大柄苹是有毒的，也很难搞清楚应该如何处理这种毒素。同时知晓这两方面知识的可能性极低，几乎可以忽略不计。然而，一旦发现了这些创新，它们便会长期有用。考古学研究表明，在澳大利亚，人们食用大柄苹的历史至少有3万年。[32]编织渔网、揉搓绳子、制作回力标和打火棒等行为，亦复如是。正是掌握了这些复杂的知识，扬德鲁万塔人才会认为库珀河流域是一片丰裕之地。就这些技能而言，掌握起来都很费力，而且其实用性也会随着人口迁移和气候变化而逐步降低。

通过其他模型所得出的结论也是如此。上文中的模型只包含了两种没有关联的特性，但实际上，很多特性都会不断变化。例如弓弩的长度、宽度、弓面、箭头、反曲度等都会发生变化，从而影响其实际效用。对于这种情况，需要假设特性的价值会不断变化，从而建立另一个模型。[33]在每一种环境中都有一个最优设计，也就是长度、宽度等的最优组合，但是在不同的环境中的最优设计各不相同。在北美东部的森林中，圆形弓面的长弓效果最好；在西部，扁平面的弓效果更胜一筹。同样，个体拥有两种不同来源的信息：来自环境中的带有杂音的信息所提示的最优特性值；从长辈等榜样身上所学到的特性值。学习就是把这两种信息融合起来。

据贝叶斯学习模型显示，最优标准即是两种信息的加权平均值，进化问题说到底就是这两种信息的权重问题。当环境信息中充满杂音，且环境变化速度较慢时，自然选择会赋予“他人行为”很高的权重，也就是说，进化的结果是，大部分个体的特性值与其模仿对象的平均值很接近。人们的学习主要以模仿为主，只有存在明确动机的情况下，才会做出实质性的改变。技术史专家已经清楚地证实了，逐步改进是如何变得多样化，如何改进工具和其他人工制品的。[34]新发现和新发明往往得自偶然，是不同技术传统要素的重新组合，而非天才的洞见，因为天才思考时能从问题一直想到答案。[35]

模仿并没有那么盲目

到目前为止，我一直都在假设模仿是盲目的。社会学习者观察长辈中某些个体的行为，然后计算出某种行为的支持者和不支持者分别有多少。换言之，社会学习者并不关心学习对象的个人特征，比如年龄、性别、声望，以及是否诚实，是否善于操控，等等。我做出这样的假设是为了让事情看上去简单些，但是从切身经验来看，这种假设是不现实的。[36]社会学习者会根据相关线索判定谁是最佳学习对象。确定榜样，意味着个体认定某些线索能够提高学习效率，而且无须理解榜样为何要那样做。现在来看看下面两个例子。[37]

假设年轻的学习者可以观察到长者的成功。一位年轻女性在学习如何采集果实时会留意哪位女性带回来的根茎最多，谁的劳累程度最低，谁的速度最快，谁是其他妇女最羡慕的对象，等等。这样一来，模仿成功者的倾向就会被这些与采集果实相关的特性得以在人群中广泛传播，即便模仿者并不清楚背后的因果关系。在可比较的特性范围较狭窄时，这一点会表现得尤为明显。你观察到姑姑善于挖掘根茎，注意到她的挖掘工具又细又长，且由较硬的木头制成。于是，你模仿她的行为，开始利用这三个特性，即便只有工具的长度真正与成功有关。随着时间的推移，长度和效率之间的关联性有了越来越多的可信数据做支撑，在这种情况下，非随机的模仿就会变成优秀的策略。对因果关系的理解是很有用的，它可以帮助个体排除不相关的特性，比如手柄上的装饰等，而且这种理解并不需要特别精准。复制不相关的特性，比如厚度或者装饰，将让这一过程变得杂乱。

社会学习者还必须避免受到学习对象的欺骗。许多特性都具有欺骗性。一个男人如果每天都在用弓箭射猴子，很有可能觉得弓箭比吹箭筒更好用。人类的社会学习心理机制很有可能是由自然选择缔造的，所以才会关注这些“能增进可信度的行为展示”。[38]

请注意，这些关于社会学习心理机制的分析是建立在进化推理基础之上的。人们应该向那些德高望重、言行一致之人学习，因为这样的行为曾为他们带来了收益。在过去十年间，越来越多的证据表明，我们的假设与其他类似推理所得出的论断是正确的。[39]

难以理解的因果关系

青少年总想说服父母让自己做一些其他孩子正在做的事。如果他们的父母如我一般，那么这些孩子可能会得到一个令人崩溃的反应：“哦，如果有人从桥上跳下去，那么你也会照做？”很难相信，人类的技术成就来自这样一个过程。难道人们不需要调整技巧以应对各种意外？在传统文化中，制作弓箭的技巧和桑橙木的特性有关。每一块木材都是不一样的：大小不一、密度各异、有无枝节等。为了制作出一张标准的弓，匠人们必须适应这些差异。难道不需要理解弓箭制作的基本原理吗？

这是很常见的一种直觉反应，但我觉得这种反应是错误的。很多动物都能制造出复杂的东西，而且是在对自身所从事的活动无甚理解的情况下。鸟巢、蜘蛛网、白蚁丘，以及河狸坝，这些由动物制造出来的事物早已为人熟知。倘若浏览一下动物学方面的文献，你还会发现一长串不那么熟悉的“动物制造”。[40]这些事物大都设计精巧、技术精湛，想一想织巢鸟织造的巢穴就知道了。

虽然这些巢穴都织造得十分精巧，但无论是观察结果还是实验研究都证实，巢穴的织造过程并非基于对因果关系的理解。相反，它更接近这样一种算法：将重复的、简单的模式化行为按照一定的顺序组合起来，一步步筑起巢穴。[41]首先，鸟儿会织出一个圆环，以便在织造孵卵室时在上面休息；然后，它会织造一个入口，如图1-6。织造算法的进化使个体鸟儿能够调整自己的行为以适应可获得的材料，以及鸟巢所在树枝的形状。

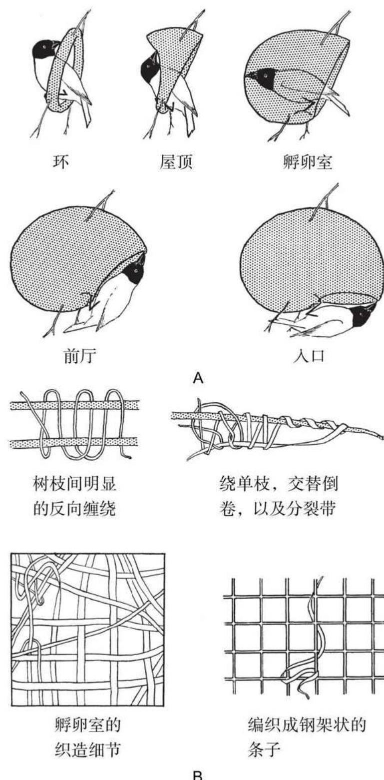


图1-6　图**A**为织巢鸟中的乡村织巢鸟织造巢穴的步骤。这种鸟儿会先将绿草茎连接在一起，织出一个悬挂在树枝上的圆环，然后将更多的草茎连接起以填补圆环，接着在既有的圆环结构上不断编入更多的草茎，将圆环扩展成一个巢穴。图**B**显示的是乡村织巢鸟的鸟巢内部结构，以及织造方法。

From Collias & Collias, 1964.

白蚁、漏斗蜂、蜘蛛之类的无脊椎动物，虽然能够制造出非常复杂且功能完备的事物，却没有相应的神经机制来构思所制造事物的最终样式。[42]事实上，甚至有些没有神经系统的生物也能制造出复杂的事物，如图1-7所示，单细胞变形虫冠砂壳虫（difflugia　corona）也能建造“房屋”。动物们在制造事物时还会受到各种环境因素的影响，比如，为了调节温度，澳大利亚北部的磁性白蚁会将蚁墩大体建造为扁平状，以及南北朝向（如图1-8），不过具体朝向会根据风向和遮挡物的位置发生变化。[43]数百万白蚁在漆黑的地下协同工作，依靠简单的算法，根据不同的条件对白蚁墩的建造做出调整。

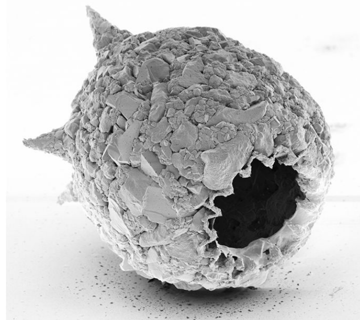


图1-7　单细胞变形虫冠砂壳虫建造的“房屋”。直径约为0.15毫米，由细小的沙子黏合而成。在变形虫分裂繁殖时，其后代中的雌性细胞将会继承这个“房屋”，另一个细胞则会建造一个新的“房屋”。冠砂壳虫分布甚广，其不少亚类也都会建造类似“房屋”，但“样式”略有不同，因为要适应当地的环境。

From Hansell, 2005.



图1-8　澳大利亚北部的磁性白蚁墩呈扁平状，其长轴通常为南北朝向，不过准确的朝向取决于风向和遮挡物的位置。

Photo by Neil Liddle.

当然，如果环境变化超出了进化算法所能涵盖的范围，那么运算就会出现极大的错误。漏斗蜂的筑巢过程就证明了这一点。[44]这些蜂会在倾斜的河岸地表之下筑造蜂巢，并在巢中孵化幼蜂。通往蜂巢地道的入口呈倒漏斗型，然后是一条长长的通往主巢的地道。地道的长度必须适中，这样才不会露出地面。漏斗蜂在“运算”时依据的是筑造时间，在筑造地道一段时间后，开始筑造漏斗。在实验中，我们抬高了地面，但漏斗蜂并没有意识到这一点，即便漏斗露出了地面也不知道。在这种情况下，胡蜂一类的寄生蜂就可以轻松地进入漏斗蜂的巢穴，并将卵寄生在里面。

同样的道理也适用于人类与人造物品。就房屋和弓箭这类复杂的人造物品而言，其背后的因果关系是比较难理解的。要想达到理想的制造效果，个体就需要与多种因素进行复杂的交互。因此，模仿他人的做法往往是一个不错的策略，即便不理解他人为什么要那么做。不过，如果环境发生了变化，或者违背了相应的因果关系，模仿也会带来不良后果。当然，在大部分情况下，模仿都会带来比较好的效果。

文化的进化速度比基因更快

啮齿目动物几乎能适应任何类型的栖息环境，其“成功”靠的是最常见的基因进化和个体学习的结合。与人类相比，文化适应性在这类物种身上表现得微乎其微。选择食物、处理毒素、避开捕食者，以及建造像水獭坝那样的事物，这一切足以与几千年前的任何一个人造事物相媲美。最大的不同是啮齿类动物有2 200个亚类，而人类没有亚种。这里的观点是：人类可以适应多种环境，而其他动物却不行，因为与基因进化相比，在积累适应性知识方面，文化进化的速度更快。人类在进入新的环境后，只需几个世纪，而非上万年，就能掌握在当地生存所必需的知识和工具，并建立起社会秩序。

这个观点是正确的吗？文化进化真的比基因进化更快？毫无疑问，在现代社会，文化变迁的速度是惊人的。我父亲是在纽约北部地区长大的，在他小时候，世界上还没有电话、汽车和电灯；而现在，他的孙辈们却能够轻松地使用着功能强大的计算机，或者坐飞机跨越大洲大洋。这种惊人的变化是过去1 000年间技术变革呈指数级发展的必然结果。[45]对于像人类这种寿命较长的物种而言，文化上的变化比基因上的变化快得不是一点点。

我在亚利桑那州立大学的同事——社会学家查尔斯·佩罗（Charles Perreault）的一篇文章有力地证明了，很长一段时间以来，文化进化的速度比基因进化的速度快得多。[46]佩罗的灵感来自生物学家菲利普·金里奇（Philip Gingrich）的一篇经典文献，[47]在那篇文献中，金里奇收集了大量古生物学方面的数据，并对古生物随着时间推移所表现出来的形态上的进化进行了测算。在金里奇的数据集中，每个点都代表着：在两个不同的时间点上，同一种生物在某些形态上的变化，比如股骨长度的变化。然后，金里奇计算出两次测量之间的变化百分比，并除以时间差。

佩罗从考古记录中收集了573个样本，主要是一些生活在全新世时期的北美地区，采集社会与农牧社会相融合的种群。每一地点都有一些人造物品，比如陶盆。通过对陶盆直径进行测量和对比，他发现这些人造物品属于同一个考古遗址在不同时期的产物。佩罗将测量到的人造物品的变化速度与金里奇样本中的古生物的变化速度进行了对比。在其他条件不变的情况下，陶盆、墙壁和房屋等事物的变化速度是下颚骨、臼齿和股骨等生理构造变化速度的50倍左右。他还进一步验证了金里奇的发现，变化速度与测量时长呈负相关。因此，似乎可以得出这样一个比较可信的结论：文化进化的速度确实要比基因进化的速度快很多（图1-9）。

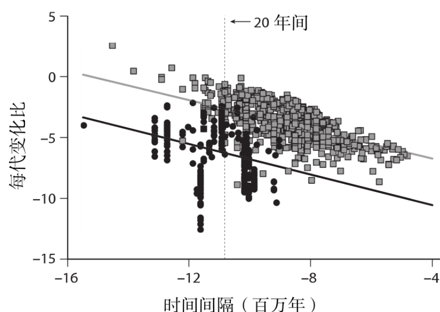


图**1-9**　在化石记录中，基因遗传的形态特性在每一代的变化百分比，如黑色圆圈所示；在考古记录中，文化传承特性在每一代的变化百分比，如灰色方块所示，除以变化时长所得到的结果。图中的实线代表多变量协方差分析（**Multivaviate Analysis of Covariance**）(3)中的最优匹配程度。在这两种情况下，时间间隔增长，变化速度放缓，而每一代的斜率基本都是相同的。两条实线之间的距离代表在控制其他变量的情况下，文化进化与生物进化的速度差异。文化进化的速度比基因进化的速度快**50**倍左右。

From Perreault, 2012.

为什么人类的适应能力如此强大

我常常在想，既然累积性的文化适应是一种如此强大的工具，那为什么只在进化过程中出现一次？这是个谜。不仅如此，在生命的发展史上，很多重大的创新也只出现一次。光合作用产生了大量的能量，释放出了大量的氧气，从而造就了这个世界的生态系统。真核细胞生物、具有大型复杂细胞的有机体，以及其他一些新品种的出现，不仅促进了所有寿命较长的大型生物的进化成为可能，也让这个世界变得更加丰富多彩。横纹肌纤维可以帮助大型动物快速移动，从而使捕食者及其猎物都能不断得以进化。这些在进化过程中都只出现过一次。[48]

英国生物化学家尼克·莱恩（Nick Lane）在其杰作《生命的跃升》（*Life Ascending*）中指出，这些创新之所以只会在进化中出现一次，是因为每一种创新既是偶然的产物，也是必然的产物。[49]一些小概率情况下才会出现的环境条件使创新得以进化，这是创新的偶然性；而创新的优势使其能够被广泛传播，则是创新的必然性。

光合作用就是一个极佳的例证。蓝藻、水藻等植物都可以进行两种不同的光合作用，一种是利用强氧化剂从水中提取电子，制造出腺嘌呤核苷三磷酸（ATP）；另一种是利用强还原剂将二氧化碳还原，生成糖和其他复杂的有机分子。其他细菌只能进行一种光合作用，有的采用第一种方式从阳光中获取能量，有的采用第二种方式从分子中获取电子，并结合二氧化碳生成更复杂的有机化合物。这两种系统是如何神奇地关联到一起的呢？按照莱恩的观点，这需要一种特殊的环境，在该环境中，紫外线辐射的强度极高，并存在大量可溶解的锰。蓝藻的祖先用锰元素来保护自己免受紫外线辐射，但这会产生自由电子，从而给第一种光合作用带来麻烦。解决之道是利用第二种光合作用将多出来的电子都吸收掉。很偶然地，某个蓝藻的祖先拥有这两种系统，而这个巧妙的解决方案促进了蓝藻和水藻等植物的进化。

那么，有哪些因素会让累积性的文化适应变为不可能呢？我们并不真的知道。或许存在这样几种可能：

低精准度的社会学习。很多有机生命体都会通过大量的社会学习来延续不同的地方性传统，但是都未能表现出累积性的文化适应这个特点。例如，某些种群中的黑猩猩会使用石头砸开坚果，而其他一些种群中的黑猩猩则不会，但是在所有的黑猩猩种群中，都没有出现更好用的砸坚果工具。结合社会线索，实际情况更有可能是这样：虽然某些行为是在模仿其他猩猩，但是不管在何种情况下，这些行为都可以通过个体学习来掌握。[50]模型推导也证明了，累积性的文化适应对社会学习的精准度非常敏感。人类学家彼得·里克森和我一道提出了一种假设：通过观察来进行学习的能力，也就是真正的模仿能力，是文化积累的基本要素之一。[51]人们一度认为，真正的模仿是人类独有的能力。然而，生物学家安德鲁·怀滕（Andrew Whiten）及其合作者们通过巧妙的实验证明，黑猩猩也可以通过观察同类来学习新的行为，只是与人类相比，其学习精准度有所欠缺。[52]因此，在文化累积方面，高精准度的模仿是必不可少的。

使用工具的潜力有限。其他灵长类动物在使用工具方面的潜力极其

有限。有的灵长类动物会使用石头砸开坚果和贝壳，黑猩猩还会使用一些其他类型的工具，但不管怎样，它们使用工具的潜力很有限，因为是四足行走的动物，所以很难将工具从一个地方带到另外一个地方。事实上，黑猩猩会在需要使用工具时就地取材，现场制作，用完就扔。显然，在制作工具这件事上，投入过多会导致得不偿失。

合作程度较低。我们会在下一章中看到，人类的合作程度比任何哺乳类动物都要高。即便是在最小规模的人类社会中，也存在各种互惠网络，而且这种网络的规模可能是猿类互动网络规模的50倍。[53]基于多方面的原因，合作会促进文化传播；其中最重要的原因是，精准的文化学习是需要教导的。归根结底，社会学习是一个推导的过程。人们并不是在单纯地模仿他人的行为，实际上是先观察他人的行为，再推导行为背后的观念。有时候这种推导很容易。如果我看到你正在吃东西，我可以确信，你认为这些东西没有毒。然而，许多类型的社会学习都会遭遇“框架问题”，也就是说，学习者需要先判断所看到的画面中，哪一部分才是值得学习的。假如我看到你在制作弓箭，而这一过程会涉及很多行为：你盘腿而坐，削尖箭头，选择好羽毛并粘在箭尾，同时嘴里反复哼唱着小曲。很难确定对于制作弓箭来说，哪些行为是必要的，哪些是不相关的。心理学家捷尔吉·盖尔盖伊（Gyorgy Gergely）和盖尔盖伊·奇布劳（Gergely Csibra）认为，除非被模仿者提供了一些线索，且足以让模仿者推断出其行为中哪些部分与需要归纳的知识有关，哪些部分无关，否则社会学习不可能非常准确。[54]

教导是一种合作性行为，对学习者有益，但需要教导者付出成本。在实践中，只是向父母学习不算是进化上的奇迹，因为培养后代能够给父母带来基因上的“收益”。但是，在传统社会中，人们会从群体成员身上学到很多知识，而且只有人类表现出了这种非亲缘性的合作。

口头指导常常会促进文化学习，因此文化学习离不开合作。例如，一项针对大学生的实验表明，在学习制作类似于250万年前的简陋的石器工具时，口头指导必不可少。[55]与那些接受了口头指导的大学生相比，那些只能依靠观察和模仿，或者只得到了一个参考物的大学生，任务都完成得很不好。

语言本身就是一个谜。在大部分物种中，发出沟通性的信号必然需要消耗一定的成本以确保其真实性。比如狮子的咆哮和马鹿的鸣叫，要想发出这些信号，这些动物就需要付出较高的成本，因此这些信号不可能是虚假信号。一头孱弱的狮子不可能发出可怕的咆哮，但是人类的语

言却“可以低成本交谈”，因此也带有一定的欺骗性。我可能会告诉别人，我在外出打猎时遇上了一头狮子，而且差点杀死它，然而实际情况可能是我躲在灌木丛中瑟瑟发抖。双向的监督和惩罚可以控制欺骗，但这要求个体之间存在合作。[56]在如今仅存的那些狩猎-采集部落中，说同种语言的人大约为1 000人，没有别的脊椎动物能与非亲缘同类达成如此大规模的合作。

学习不到什么有用的东西。我们有充分的理由相信，那些使累积性文化学习成为可能的心理能力或许最初并不具有适应性；直到后来，当它们广泛存在于在人群中时，才具有了适应性。[57]我们假设，对于累积性的文化进化来说，新的认知机制是必要条件。比如，很多学者都认为，精准的文化学习需要人类能够对他人的信念进行思维上的表达，而这种能力被称为“心理理论”（theory of mind）。

例如，我看到了一条蛇，同时灌木丛挡住了你的视线，所以你看不到蛇。我如果不懂这种心理理论，就不会意识到你并不知道那里有一条蛇，因此，当你径直走向那条蛇时，我会错误地以为你认为这条蛇无毒；我如果了解这种心理理论，便会意识到你没有看到那条蛇，因此会知道，你的行为基础并不是“蛇是否有危险”这一认知判断。倘若这种推理是正确的，那么对于文化进化而言，心理理论就是必不可少的。

黑猩猩似乎对同类的所见、所闻和所想都有一定的了解，但其心智能力不如人类这么强大。假设远古人类的心智能力与黑猩猩差不多，在这种情况下，极少数心智能力超常的个体可能并不会受到自然选择的青睐，因为整个群体并不具有累积“地方适应性知识”所必需的认知机制。如果这一假设是正确的，那么心智能力可能就与其他一些因素有关，比如马基雅维利式的操控他人的需要(4)。这种能力一旦出现，累积性的文化传播就会开始。

过往的“文化幽灵”

我在这里对文化进化所做的阐述，假设了历史会对文化适应产生影响，就像影响基因适应一样。达尔文在剑桥大学读书时由衷地崇拜威廉·佩利（William Paley）神父，并十分迷恋其著作《自然神论》（*Natural Theology*）。在这本书中，佩利热情地歌颂着自然界中的各类精妙设计，比如，极为详细地描述了眼睛的结构特点与视觉功能的关系。在佩利看来，眼睛和其他类似的复杂事物皆出自一位无所不知、无所不能

的“设计师”之手。今人大多都不会相信这一观点，但在19世纪，这一观点流传甚广。这便带来了一个有趣的话题。在维多利亚时代，许多人对适应性本身并无质疑，他们所质疑的是，在适应过程中产生的那些怪诞的“非优化设计”及其残留物。既然能够创造出眼睛这样的一流“光学仪器”，为什么“设计师”还会草率地拼凑出人类脊柱这类粗制滥造的东西？为什么蝙蝠的翼骨会更接近鼹鼠或儒艮的四肢，而非鸟儿的翅骨？

当然，今天的人们自然知道，进化的基本过程是这样的：不断地在原有的基础上缓慢地改进，从而发展出新的适应能力。这便意味着，如今某一血统所青睐的特性类型取决于其生活在遥远过去的祖先的适应情况。蝙蝠的翼骨由手掌进化而来，而鸟类的翅骨进化自恐龙的前肢，这是很不一样的。人类的脊椎最初是为了适应四足行走而出现的，后来为了适应直立行走而被不断改进，最终成了现在的样子。这并非像你在一张白纸上作画那样随意，一下笔就是一块椎骨。

历史在很大程度上也影响着文化适应。因为文化适应是一个渐进的过程，所以有些行为变异可以用种群的历史来解释。在那些相对独立且存在时间很短的社会中，人们在行为和观念上的相似度非常高；在那些相对独立但存在时间很长的社会中，人们在行为和观念上的相似度则非常低。另外，对于相同的环境问题，不同的种群也会进化出不同的解决方法。然而，我们并不清楚这些来自历史的影响有多重要。一方面，许多人认为，人类的智慧足以解释为何人们能够适应不断变化的各种环境。倘若真是如此，那么便可以确信，人类能够快速地适应新环境，过往的“文化幽灵”很快就会被征服。另一方面，如果你跟我一样，认为文化适应的累积需要经历千秋万代的话，那么显然，文化变迁势必要经历更长的时间，而且过往的“文化幽灵”则会不断地产生影响。

很遗憾，我还没来得及着手整理最近的一项研究结果。这项研究针对的是北美西部土著社群的文化变迁模式。该研究由我所在的人类起源研究所（Institute of Human Origins，IHO）的同事萨拉·马修和查尔斯·佩罗主持。[58]研究基于北美西部印第安部落的相关数据，该数据由加利福尼亚大学伯克利分校的人类学家A. L.克罗伯（A. L. Kroeber）带领学生于20世纪初期收集。学生们被分派到北美西部的各地，对当地老人进行走访，以了解在欧洲人到来之前当地人的生活状况。克罗伯非常明智地从不同种群中收集到了同类信息。这些资料包含了与172个种群的272种行为特质有关的信息，这些种群的生活范围非常广泛，覆盖了北美西部的大部分地区。在这些种群中，绝大部分部落都属于小规模社群，而

且许多都还处于觅食经济阶段。这些被记录下来的行为特质包括：技术特征，比如是否会编制筐篮；生存之道，比如是否会种植玉米；反映社会组织形式的特质，比如是否存在母系氏族，以及一些仪式方面的特征，比如逝者是否会被埋葬；等等。

在克罗伯去世之后，人类学家哈罗德·德赖弗（Harold Driver）及其学生约瑟夫·乔根森（Joseph Jorgensen）对这些资料进行了整理，并补充了每一部落所处环境的气候资料，以及100种以上的动植物的诞生与消亡资料。萨拉·马修和查尔斯·佩罗将这些行为学和生态学上的特征转换为457个二分类变量，同时还增加了一些资料，补充说明了不同部落之间在语言学和地理学方面的共通性。语言学方面的共通性代表的是共同的文化历史，而地理学方面的共通性则代表着文化特质的传播所带来的相邻部落之间的文化相似性。

马修和佩罗利用这些资料比较了文化史与生态环境在文化特质的变化模式方面所起的不同作用，[59]并对不同时期文化特质的变化速率进行了考察。在图1-10中，他们将行为特征分成若干类别，并总结了文化史与生态环境的效应比值（ratios of the effects）。被用来做比较的指标是经过标准化检测的，因此当数值大于1个平均值时，文化史所解释的变异量远远超出生态环境所解释的变异量；而当数值小于1个平均值时，生态环境的作用更重要。可以看出，除生存特征之外，在各类行为特征上，文化史的作用都比生态环境重要得多；而在生存特征上，文化史和生态环境的作用不分伯仲。这意味着，如果你想知道一个人是否会编制篮子，一位女性是否属于某个母系氏族，某个部落是否会将逝者埋葬，那么知道他们所讲的语言要比知道其所居住的环境更为有用。令人惊奇的是，无论是语言还是生活，都可以很好地帮助我们判断出一个人是否会种植玉米。过往的“文化幽灵”会对人类生活的诸多方面产生影响。

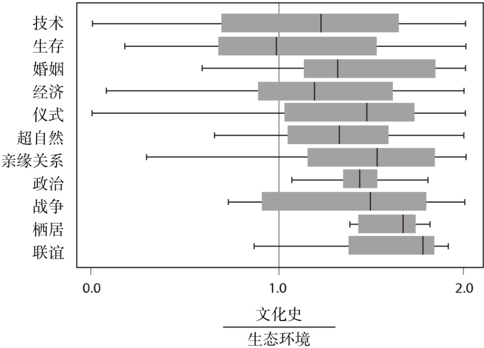


图1-10 条形表示的是在解释不同类型的特质时，文化史与生态环境的相对重要性。对每一种类型来说，条形代表的是相对重要性的数值范围。数值大于1，意味着文化史更重要；数值小于1，意味着生态环境更重要。每个条形中的垂直线表示的是该类特质的中值。条形代表的数学区间是第25到第75的百分位数，误差是四分位数间距的1.5倍。除了“生存特征”这一类型，在其他所有类型中，文化史对特性值的预测都要比生态环境更准确。

因此，过往的“文化幽灵”会存在相当长的时间。为了弄明白这一问题，马修和佩罗将高级语言分类作为预测变量，重新进行了分析。要理解这里所说的高级语言分类，可以联想一下英语及其相关语系。荷兰人和德国西北部的人们所讲的弗里斯兰语（Frislan），与英语的关系最为密切。这两种语言都源自撒克逊语（Saxon），而撒克逊语早在1 000多年前就盛行于环北部海区。相应地，它们可以被划分到更高一级的类别中，也就是盎格鲁-弗里西语（Anglo-Frisian），而这一语系的语言又同德语以及荷兰语的关系最为密切。德语和荷兰语又都源自一个更为古老的语言社区内所讲的语言。

以上四种语言又都被归为更高级别的西日耳曼语（West Germanic）。西日耳曼语又同斯堪的纳维亚语（Scandinavian）融合，形成了日耳曼语。最终，日耳曼语又同其他一些欧洲语言，如罗曼语（Romance）、斯拉夫语（Balto-Slavic）、亚美尼亚语（Amninian），以及包括波斯语（Farsi）和印地语（Hindi）在内的一部分印度-伊朗语（Indo-Iranian）一同被划分到了印欧语系的语门。印欧语系中的语言具有一些共同特征，这些特征源自大约7 000年前在安纳托利亚（Anatolia）地区盛行的一种古老语言。高级类别对于行为的预测能力远弱于低级类别，因为语言的高级类别所涉及的文化联系通常都比较深远。这正是马修和佩罗所发现的研究结果，如图1-11所示。

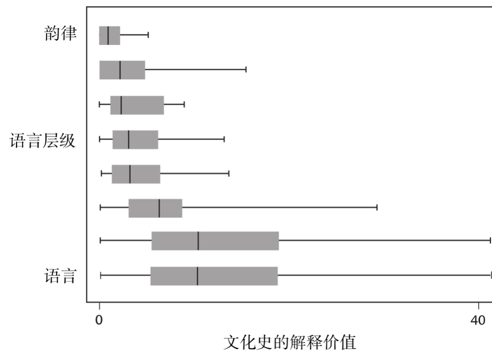


图1-11 在解释使用不同层级语言的人群的行为变化时，文化史所起的重要作用。最低级的类型是个体语言，最高级的是语门。预测能力下降得最明显的是第二级和第四级的语言。这些语言大体上相当于印欧语系中的罗曼语和日耳曼语。这意味着文化史的变化是以千年为时间刻度的。

Redrawn from Mathew & Perreault, 2015.

然而，值得注意的是，在第二级与第三级之间、第三级与第四级之间，预测能力皆出现了明显的下降。这些语言都属于萨利希语（Salishan）和瓦卡希语（Wakashan）之类的语系，大体上相当于印欧语系里的日耳曼语和罗曼语。例如，人们会认为，各类萨利希语都源自同一种古老的语言，这种语言出现的时间在1 000年以上。这就意味着，即便知道某人的祖先在千年之前所讲的语言，也很难以此来了解他今天的行为特征。

这一研究的结果与大量文化系统发生学（cultural phylogenetics）的研究结果是一致的。在过去10年左右的时间里，文化系统发生学非常盛行，它旨在对传统滑雪之类的具体技术特征[60]，以及诸如政治等级[61]、婚姻制度和生育行为等社会特征的发展历史进行重构。[62]该研究表明，除了极少数情况之外，在语言上存在密切关系的种群，在其他方面也具有很多相似之处。这也进一步验证了：文化史会随着时间的推移而进化。

人口规模与文化适应

还有一种方法可以证明，人类的适应性是基于缓慢的文化累积的。假设其他生物都和人类一样，与那些个体数量较少且与世隔绝的种群相比，个体数量多的种群拥有更多样且更复杂的工具。这一假设可以用两种不同的观点来进行解释。

第一种观点是：接受新变异的个体数量会受到一些偶然因素的影响。[63]这就意味着，如果这些新变异的创造者没有被他人模仿，那么新变异就有可能会随机遗失。例如，在一位聪慧的妇女不幸意外死亡之后，她所拥有的食物烹制知识就有可能随之遗失。小规模部落中的这种由“文化漂移”（cultral drift）所致的新变异遗失，在概率上往往会高于大规模部落，因为在小规模部落中随机遗失更有可能发生。

第二种观点是：社会学习常会遭遇一些错误，而这些错误会弱化那些复杂的适应性特质所起的作用，所以大部分“学生”都达不到其“老师”的水平。个体拥有选择性学习的能力，会向实践经验丰富的专家学习，而这种能力可以防止文化遗失现象的发生。如果有极不寻常的“学生”超越了其“老师”，那么累积性的文化适应就出现了。约瑟夫·亨里奇

将这一过程比喻为：跑步机上的跑步者需要保持好姿势。[64]大规模群体中的学习者有机会接触到更多的专家，因此他们的能力也更有可能获得提升。

在上述两种过程中，遗失的文化适应的多样性可以通过人与人之间的接触而得到补充，进而产生更高水平的变异，以及出现更具适应性的特质。[65]

我和西蒙弗雷泽大学（Simon Fraser University）的人类学家米歇尔·克兰（Michelle Kline）做了一项研究：人口规模和人际接触对大洋洲岛屿种群的海洋觅食工具包的影响，以及影响的复杂程度。[66]这些工具包主要包括渔网、鱼叉和鱼钩等。我们之所以选择岛屿种群进行研究，是因为其拥有明确的地理边界，也就能对人口规模和接触频率做出比较准确的评估。在调查中，各种群所使用的海洋生态系统都非常相似，可以最大限度地降低对生态多样性的影响。这些种群还拥有共同的文化传统，这就减少了对文化史的潜在影响。研究结果表明，工具的总类型数以及每一类工具的复杂程度，都与人口数量呈显著的正相关关系。个体间存在更多交往的社会，要比那些个体交往少的社会拥有更多更复杂的工具。见图1-12。我们还考察了其他一系列预测变量，但那些变量都无法对各岛屿社会的变异量做出解释。

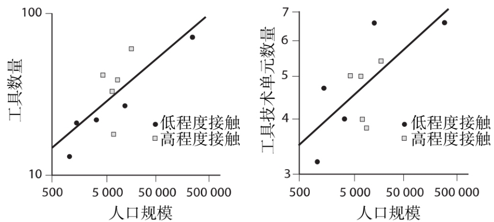


图1-12　左图表示的是，不同的太平洋岛屿社会所拥有的海洋捕捞工具的数量与人群规模的比值。右图表示的是，每件工具所包含的技术单元数量与人口规模的函数关系。技术单元数量可以被看作技术复杂性的粗略指标之一。所使用的是对数刻度，因为文化漂移模型和文化跑步机模型都认为，在人口规模和文化复杂性之间存在着一种凹形关系。所使用的数据来自“人类关系区域档案”（**Human Relations Area Files**）世界文化数据库（**World Cultures**）中的人种志电子信息。数据库的编码者对海洋捕捞工具的数量和技术复杂性的平均值进行了估算。对“较低程度接触”与“较高程度接触”的认定，主要依据“人类关系区域档案”中的数据编码。我们的研究对象包括了所有的大洋洲群体社会，因为采用了多重编码的形式，所以这些人并不知道这个需要被验证的假设。

From Kline & Boyd, 2008.

社会学习是否会促进复杂行为的累积性进化

我们还可以用模拟实验来对“累积性的文化进化”模型进行验证。我曾试图向一位生物学家朋友描述我在这方面的研究，但他立即问道：“是什么物种？”这里出现了误解，我所说的“实验室文化”并非是指在实验室器皿中培育出来的文化，而是指实验者在社会中进行传播的文化。某个人一旦学会了某件事情，就会成为群体中一个新的信息来源，继而该群体也会成为其他群体的信息来源。利用这类实验可以系统地记录下，在规模较小的群体中信息是如何传播的。

我与人类起源研究所的同事马克西姆·德雷（Maxime Derex）使用下面这个方法检验了：社会学习是否会促进复杂行为的累积性进化。[67]在实验中，参与者被要求创造出类似于图1-13C所示的图腾柱，并将因此而获得报酬。一开始，参与者被允许使用6个最基本的材料，如图1-13A所示，并可以通过专门的小组研讨会将这些基本材料合成为高级材料，最终制成复杂的工具，比如斧子或绳子，如图1-13B。参与者可以利用这些工具来砍树、制作圆木，以及制作图腾柱，如图1-13C。另一些高水平的创新，比如雕刻工具或者颜料，可以用来美化图腾柱，从而提升其价值。参与者的分数主要取决于其发明数目，以及所制作的最好的图腾柱的价值。在图1-13C中，左边的图腾柱需要22项发明，可获得920分；右边的图腾柱需要54项发明，可获得6 526分。

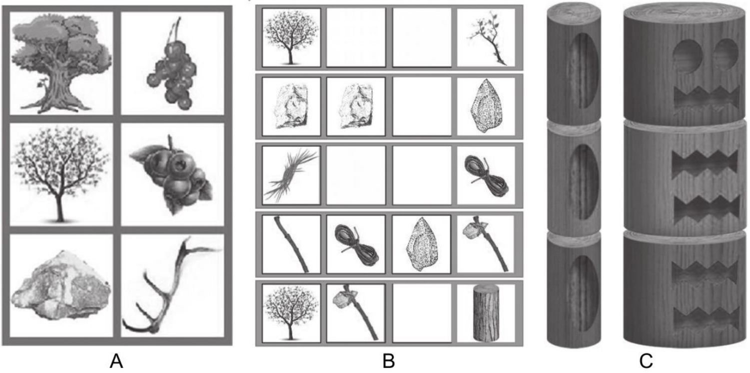


图1-13　实验中所采用的计算机任务的屏幕截图。

From Derex & Boyd, 2015.

为了对比个体学习和社会学习的效果，我们改变了实验参与者的可用信息。在个体学习处理组中，参与者只能获得与自己的发明有关的信息，而后我们将这种条件下的结果，与另外几种社会学习条件下的结果进行了比较。在部分信息处理条件下，每个参与者可以看到其他5个参与者的发明，但无法看到这些发明的产生过程。例如，一个人了解到另

一个人发明了斧子，但并不了解制作斧子所需的较低级别的发明，以及这些发明的组合方式。在完整信息处理条件下，每个参与者都可以看到其他人的发明，以及发明的产生过程。

该研究的结果表明，社会信息可以极大地提高人们解决复杂问题的能力，同时其属性也是非常关键的，如图1-14。完整信息处理组中的大部分参与者的得分，都高于个体学习条件下表现最佳者的得分。部分信息处理组中的参与者所获得的中间分数（intermediate score）表明，通过个体学习是可以制造出某种特定工具的。当然，该实验无法说明个体学习者是如何制造这种工具的。

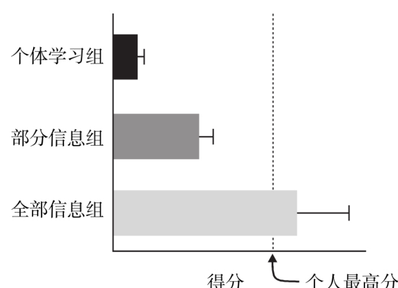


图1-14 条块分别表示个体学习组、部分信息组、全部信息组的平均分，误差条表示标准误差，短线表示个体学习组的最高分，该值小于完全信息组的平均分。

From Derex & Boyd, 2015.

文化的适应性

希望你们已经相信，人类拥有向他人学习的能力，而且这种能力可以确保人类群体拥有超强的、针对局部环境的文化适应性。即便是在最简单的狩猎-采集社会中，适应环境所需要的工具和知识也是相当复杂的，单靠个人学习不可能全部掌握。

然而，一些读者可能会想："人们应该不会从适应主义的视角去看待那些无足轻重之事吧？"每年都会有登山队向珠穆朗玛峰发起挑战，尽管他们知道，平均每20人就可能会有1人无法生还。博士生们为了获得博士学位会推迟生育计划，而后又会为了获得终身职位进一步推迟生育时间，尽管他们知道，获得相应职位的机会非常渺茫。人们会对人口数量以百万计的种族产生认同，还会青睐圈内人而排斥圈外人。类似的行为数不胜数。像老鼠和狒狒这样的动物压根不会考虑文化问题，也根本不会表现出上述行为；其行事原则就是看实不实用。而人类虽然拥有发达

的大脑和强大的推理能力，却会干出许许多多的蠢事，这对人类的繁衍生息是极为不利的。

以列达·克斯米兹（Leda Cosmides）和约翰·图比为代表的研究者们从进化的角度研究心理学，并提出了一种见解。在他们看来，人类的各种奇怪行为可以用一句话来概括，那就是“当今时代的脑袋，石器时代的心智”。[68]这些研究者认为，人类的行为源自一种进化的心理状态，这种心理状态会在极其有限的范围内影响选择性行为，比如如何择偶、是否重视社会伙伴、如何看待工具等。这种心态背后涉及的心理机制可以帮助我们对特定的环境线索做出有效的反应。例如，对异性的潜在吸引力会受到年龄的影响。

这些反应所适应的环境正是其进化时所处的环境，一般认为是中更新世到晚更新世时期的狩猎-采集社会。当时，人类的活动范围较小，只拥有简单的技术，社会关系也仅限于面对面的交往。粮食的生产，尤其是畜牧业和农业给人类带来了全新的环境，而这种环境阻断了线索与结果之间的关联，因此，这种进化而来的心理状态导致了现代人行为上的非适应性。正如我们将在下一章中看到的那样，自然选择更青睐建立在亲缘关系上的大规模合作。人类当前的大规模合作都是建立在相对独立的个体基础之上的，这与其他哺乳类动物完全不同。一般的解释是，这种奇特行为背后的心理机制进化自以亲缘关系为基础的小规模社会。

我对文化进化所做的描述，提供了关于人类行为独特性的解释。文化适应性总是伴随着一种内生性的权衡，有助于人们获得复杂的、难以习得的环境适应能力，这种能力单靠个体学习是无法掌握的。但是，只有当个体愿意接受周围人的行为时，这种文化适应性才会起作用，也就是使他人的行为和自身的经验相冲突。这种模仿他人行为的倾向有助于个体获得任何一种大众化行为，只要这种行为与个体自身的判断不存在明显的冲突。这就意味着，如果存在一些认知过程或者社会化过程，能使一些非适应性的观念为人所知并被人接受，那么这些非适应性的观念就能得以传播。[69]

进化心理学家认为，自然选择造就了人类的心理机制，确保人们做出适应性的行为反应。这又意味着，进化而来的心理机制应该能让人们成为思想市场上“谨慎的购买者”。那些与人类大脑的进化环境相匹配的观念，我们都应该接受。我认为这是正确的。总体来讲，这也是文化进化之所以能推动适应性行为进化的主要原因。

然而，学习机制并非在任何情况下都会偏好适应性行为。这种大体

有用的学习机制所导致的结果，有时候完全不同于以适应性为基础所做出的预测。例如，细微的认知偏差对人类个体所做决策的影响微乎其微，但是经由文化传播，在代际间被不断重复后，这种影响将变得异常巨大。

关于文化学习的实验验证了这一点。加利福尼亚大学伯克利分校的认知科学家汤姆·格里菲思（Tom Griffiths）及其合作者在实验中向参与者呈现了50对数字，然后要求他们寻找数值*x*与*y*之间的关系和规律。例如，我向你呈现以下数对：2和4、5和10、33和66，你会推断出$y=2x$。在其中一个实验中，*x*-*y*数对是从抛物线上选取的，如图1-15所示。格里菲思及其合作者用第一位参与者的习得性行为来训练第二位参与者，再用第二位参与者的习得性行为来训练第三位参与者，以此类推。心理学家将这种结构称作“传动链”。该实验的结果见图1-16。[70]

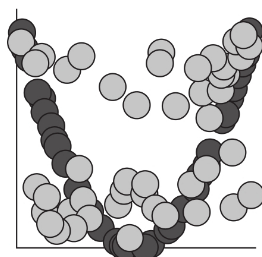


图**1-15**　黑点表示格里菲思及其同事所做实验中的**x-y**训练集。在看过训练集之后，要求一次只看一对**x-y**；实验人员会给参与者**15**个**x**值，让他们依据训练集中的规律推算出**15**个**y**值。灰点表示参与者的实验结果。这是个比较困难的实验任务，但是实验结果并没有明显的偏差。

From Beppu & Griffiths, 2009.

如图1-16所示，最后的结果与最初的抛物线大相径庭。这表明，刚开始*x*和*y*的关系并不重要。通常，传动链会在最后呈现出正向的线性关系，因为人们都有细微的认知偏差，也就是倾向于将两个数字的关系看作线性关系。这也是为什么你会觉得我的例子很简单。在实验室之外，即使是非常细微的认知偏差也可能会导致极其严重的影响。不少研究者认为，超自然的观念就是在类似的过程中出现的。因为人们会有认知偏差，总是倾向于接受那些基于主体行为的解释，因此在经过重复的文化传递之后，细微的偏差被放大了，最终导致“鬼神”和“黑色魔法”等迷信思想的出现。[71]

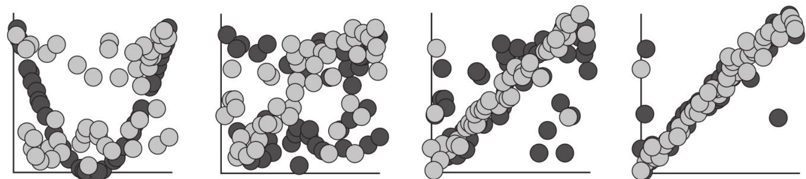


图1-16 格里菲思及其同事所进行的传动链研究结果。在第一个实验中，即最右图所示，黑点代表实验人员给出的数对关系。在后续的实验中，黑点则代表前一位参与者在其训练集中所给出的数对关系。灰点代表的是当前参与者所认为的数对关系。

From Beppu & Griffiths, 2009.

具有适应性的社会学习的认知偏差，可能会给某些特定领域带来一些非适应性的结果。我们已经了解到，个人的成功间接证明了“模仿他人是有用的”，而且数据和理论研究均表明，人们倾向于模仿那些享有一定声望的个体，这也是人类社会学习心理机制的重要组成部分。[72]然而，什么是成功？在不同的地方答案也不尽相同。比如，在一个地方，成功是参加战斗的勇气；而在另一个地方，成功是政治上的叱咤风云。这便意味着，自然选择没有为人们提供一种普适性的“名望心理”。人们必须到环境中去了解什么是声名显赫，而周围人如何获得声望无疑是一个有用的线索。

对他人的学习对象进行模仿，是一个具有普遍意义的良好规则，不过有时候可能会带来过激的反馈过程，导致声望系统的标志性特征不断被夸大。[73]这可以用来解释非适应性声望系统的文化进化过程。在这种非适应性声望系统中，人们冒着生命危险爬上雪山顶峰，或者为了在《自然》杂志上面发表文章而放弃生育。去模仿那些可以提高公信力且能够为人所见的行为，实则是一种认知偏差，可能会导致令人意想不到的副作用。比如，一些信仰系统在传播中要求信奉者从事某些高代价的公共行为，譬如“发誓终身不婚”，或者笃信“虽然贫穷，依然炫耀”。

文化适应性所造成的最令人惊讶、最违反直觉的负面影响莫过于，促使那些大规模群体中的毫无关系的成员形成某种代价高昂的合作性行为，比如战争。这类行为在其他哺乳类动物中是看不到的。人类这一物种之所以如此“卓绝不凡”，文化适应性“功不可没”。我将在第2章中就这一主题进行阐述。

02
人类的独特之处（二）：社会规范

专业化分工和交换活动

经济学家米尔顿·弗里德曼（Milton Friedman）曾用一支普通的铅笔说明了人类这一物种的超级合作性：

> 请看这支铅笔，在这个世界上，没有谁能独立制造出这支铅笔。据我所知，制造这支铅笔所用的木材来自华盛顿州一棵被砍伐的树。砍伐这棵树需要一把锯子，制造锯子需要钢铁，而炼制钢铁则需要铁矿石。黑色的铅笔芯，虽然被称作“铅”，但实际上是一种被压缩的石墨。我不确定石墨来自哪里，可能开采自南美的一些矿山。红色的顶部是橡皮擦，里面的橡胶可能来自马来西亚。至于这个铜箍、黄色的油漆，以及画出黑色线条的颜料，还有把各部分黏合到一起的胶水，我完全不知道它们来自哪里。然而我知道，通过成千上万人的合作，才有了这支铅笔。[1]

弗里德曼表现得并不是特别专业。制造铅笔所用的木材是来自加利福尼亚州的香杉树，笔芯由石墨和黏土混合而成，石墨大多来自中国和印度，金属环一般都是铝制的。[2]不过，这也正好印证他的核心观点：你不需要知道太多细节，因为神奇的市场力量能够解决所有的细节问题。

这里颇具争议的一点是，市场到底有多神奇；而关于专业化分工，以及交换活动在现代人类社会中所发挥的重要作用，则没有引起争议。我们生活的方方面面，比如衣食住行，都依赖于一个复杂的、全球性的，充满专业化分工与交换活动的网络。

从本质上来讲，专业化分工和交换活动就是合作。[3]专业化分工提高了生产和服务的效率，而交换活动则可以实现利益共享。但是，交换活动也为投机取巧提供了机会。因为除了最简单的交换活动之外，大多数交换活动都存在滞后现象。你先生产产品，我随后再给你报酬，这就意味着，我总是有机会拿走你的产品而不给你报酬。这种情况往往会出现在契约难以执行的时候，同时处于经济学中的空白地带。在洛杉矶，一些不择手段的人力资源主管会在未签署协议的情况下使用劳动力，并在几周后将劳动力无偿辞退。[4]很多交易活动都存在信息不对称的问

题。卖方对产品的了解通常要比买方多，在这种情况下，只有当买方充分信任卖方时，交易才能达成。

在规模较小的社会中，专业化分工和交换活动往往都很少。在游牧社会里，人们常常群居而生，每个群体的人数只有20到40人，即便如此，专业化分工和交换活动也是不可或缺的。深入调查发现，几乎在所有的觅食部落中都存在食物共享的现象。

在一项意义重大的研究中，人类学家希拉德·卡普兰（Hillard Kaplan）和金·希尔（Kim Hill）等人对南美的两个觅食部落进行了调查。他们统计了不同年龄段的男女对日常生计所做的贡献。[5]研究结果如图2-1所示。年轻男性和老年男性的消费多于产出，而中年男性则会产出大量的剩余价值。年轻女性的消费也多于产出，而中年女性和老年女性则一般收支相抵。这些觅食部落的经济活动主要依赖于专业化分工和交换活动。中年男性主要负责生产食物，并为其他年龄段的成员和女性，尤其是处于育龄阶段的女性提供额外的食物。中年女性主要负责照料孩子、加工食物、制作手工艺品，以及其他基本任务。

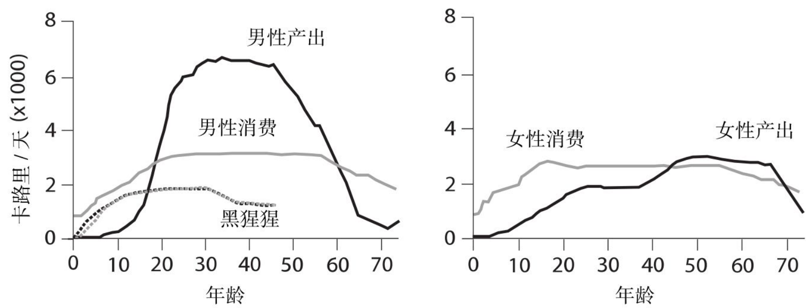


图2-1 图中所示当代的两个觅食部落：希维族（**Hiwi**）部落和阿齐族（**Aché**）部落中，两性所生产和消费的食物资源数量。纵坐标表示每日消耗的卡路里，横坐标表示年龄。左图是黑猩猩群体的数据。无论是雌性黑猩猩还是雄性黑猩猩，在断奶之后，生产的食物和消耗的卡路里几乎相等。而在人类群体中，中年男性会生产出更多的食物，并将这些食物分配给其他年龄段的成员和女性成员，因为这些人从事的是其他类型的生产活动。

Redrawn from Kaplan et al., 2000.

这种形式的食物分配和交换活动已经持续了相当长的时间，并推动了人类生活方式的进化。人类个体的生长发育速度相当缓慢，同时生育间隔期又相当短，和其他物种完全不同。这就意味着，人类需要同时照料多个孩子，而猿类母亲则不用。卡普兰、希尔及其合作者们提供了令人信服的证据，证明了专业化分工和交换活动为人类生活史拉开了序幕。

不仅如此，在觅食社会中，人们还会交换各种各样的服务。根据希尔的描述，狩猎-采集者在各种环境中都能够进行合作：

> 在阿齐族的森林聚居地，觅食之外的合作包括：为他人清理营地；为他人取水；替人收集柴火；替人点燃或照料火堆；为他人烹煮或加工食物；建造一个与他人共享的营地；制造、修理或出借各种工具；帮他人梳头；帮他人驱赶蚊虫；帮他人照料婴儿或未成年子女；喂养他人子女；教导别人的孩子；关心生病的人；为他人搜集药用植物；聆听别人的问题并给出建议；陪伴独守营地或独自外出的人；应他人要求逗人开心，如唱歌、开玩笑、讲故事。[6]

很少有别的哺乳类动物会显示出丝毫专业化分工或交换活动的迹象，尽管它们也能够从中获得巨大的收益。[7]

图2-1还描述了黑猩猩生产食物和消耗卡路里的情况。每只黑猩猩在断奶之后所消耗的卡路里和其产出是持平的。几乎不存在任何交换活动：每只成年的黑猩猩几乎只采集和加工自己的食物，而且每天晚上也只为自己搭建休息场地。尽管有时候也会分享一些肉类，但这只占其饮食的一小部分。[8]这类情况在哺乳类动物中是非常典型的。尽管哺乳类动物的合作养育方式已经进化过数次，但这种生活方式在诸如裸鼹鼠、缟獴、猫鼬、绢毛猴、绒猴、野狗、狼和海狸等物种中并不常见。[9]此外，许多进行合作养育的哺乳类动物的合作效率也没多大提升。处于支配地位的个体会压制被支配者的生育行为，被支配者别无选择，只能替支配者养育后代；就平均适应度而言，它们只能得到很少的回报，甚至得不到任何回报。

除了哺乳类动物之外，在其他许多种类的动物群体中，也存在各种各样的专业化分工和交换活动。不过在讨论这些之前，不妨先来探究一下另一种普遍存在于人类社会，但鲜见于其他大多数哺乳类动物群体中的合作行为：公共产品的生产。

这一切并不是靠你个人创建起来的

2012年，美国总统奥巴马在弗吉尼亚州的罗阿诺克市（Roanoke）发表了演讲《这一切并不是靠你个人创建起来的》（*You Didn't Build*

That）。对于福克斯新闻频道的观众而言，该演讲可谓大失人心。奥巴马似乎是受到了参议员伊丽莎白·沃伦（Elizabeth Warren）在网络上发表的一段轰动一时的即兴演讲的影响。[10]沃伦在演讲中这样说道：

> 在这个国家里，没有人能够仅凭一己之力便成为富人。绝对没有。你建了个工厂，这对你是有利的。但我想说明一下，你要把你的产品运到市场上去，而运输时驶过的公路是我们都付了钱的；你要雇用工人，而他们的教育费用也是我们支付的；你在工厂里很安全，那是因为我们每个人都为警察局和消防队支付了报酬，你不必担心劫匪们会来抢夺工厂里的任何东西。

沃伦指出了另一种关键的合作形式：公共产品的生产。虽然并不是每位成员都为公共产品的创造做出了贡献，但是诸如公路、教育、治安，以及集体防御体系之类的公共产品，为现代社会的每一位成员都带来了极大的收益。众所周知，在这种情况下，总有人想投机取巧，不付出一分一毫就坐享其成。交通、教育、集体防御体系，以及税收和法律之类的公共产品都属于惩治犯罪、砺世磨钝的现代国家制度；而另一些公共产品，比如电台等，则并不具有强制性。现代社会的生存与发展，主要有赖于公共产品的生产。

即便是在小规模的狩猎-采集社会中，公共产品也发挥着重要作用。为了改善当地环境，觅食者会通过焚烧或其他方式来进行公共建设。他们会建造一些基本的公共设施，比如车道和鱼堰，还会参加战斗。这些公共产品的生产常要涉及成百上千人，所以平均每个人所做出的贡献并不算多。

无论对于何种性质的人类社会而言，公共产品的生产都是至关重要的。这种现象在其他物种中很罕见。黑猩猩不会通过筑篱笆来隔离邻居；狒狒也没有形成共同育儿的合作行为，尽管这种合作可能对所有个体都有好处。在蜜蜂和蚂蚁的世界里，大规模的群体性合作会在创造公共产品的过程中得到进化，因为合作者之间存在亲缘关系。而人类的合作则完全不同。即便是在最简单的人类社会中，为了生产公共产品，毫无关系的群体成员也会进行合作。

在这一章中，我将证明，人类社会之所以如此，可能是因为人们的行为会受到以文化形式传播的社会规范的调控。在某种文化环境中，人们会持有相同的是非观，这些观念会推动奖惩制度的形成，从而影响人

们的行为。我们在上一章中提到过，文化适应意味着人们需要自愿接受周围人的观念。接下来将讨论，文化适应是如何激发出社会规范，以及如何促使人们生产公共产品的。在本章的最后，我还将证明，以文化形式传播的道德观念同样可以用来解释，为何对于其他动物来说专业化分工和交换活动同样不可或缺。

小规模社会中的大规模公共产品生产

在此，我准备多花一点儿的时间来讨论狩猎-采集部落中的大规模公共产品生产现象。很多研究者都认为，大规模合作是近代人类社会才出现的现象，而在更新世[11]的觅食社会中，合作规模尚小，而且受制于亲缘关系和互惠互利的动机。事实上，在觅食部落中随处可见大规模的合作行为。这一事实恰恰支持了相反的观点：在相当长的时期内，文化适应一直在促使人们成为超级合作者。

大多数觅食部落的人们都会从事各种旨在提高当地生产力的活动。他们普遍会利用烧荒来创造生产效率更高的植物群落。烧荒之后出现的植物大多是生长快速的物种，而且能承载更多的动物。[12]例如，澳大利亚西部玛杜（Mardu）部落的沙漠土著就会在冬季放火焚烧草原。

人类学家丽贝卡·布里奇-伯德（Rebecca Bliege-Bird）及其合作者通过研究证明，烧荒能够为栖居地创造更高的狩猎回报，这些回报主要是巨蜥一类的小型猎物。[13]正如这些研究者所说，烧荒所带来的环境变化或许应归类为公共产品。参与烧荒的人会付出一些成本，但受益的却是所有人。不过，烧荒属于一种低成本的活动，因此其结果很可能是，参与者会获得相应的收益，且足以抵消其付出的成本。若真如此，那么这就不算是典型的公共产品了。

当然，觅食部落的人们也会在改善生态环境方面投入巨大的人力物力。比如，居住在密西西比河与科罗拉多河周边的印第安人，会在洪水季节过去后，在裸露的泥滩上撒播野草种子；[14]另一些部落的人会移植一些块茎植物或果树；巴拉圭的阿齐族人则会将树木砍掉，待数月后返回，从枯树干上捕获大量的甲虫幼虫。[15]或许最突出的例子来自朱利安·斯图尔德（Julian Steward）所研究的加利福尼亚州欧文斯山谷（Owens Valley）中的帕尤特（Paiute）部落。[16]他们在这个山谷中修建了导流坝和运河用以灌溉土地，促进水生作物的生长，以期收获更多可以作为食物的根茎植物。最大的灌溉区占地约10平方千米，其水源主

要来自数千米长的运河。这样的建筑工程几乎需要所有当地人的通力合作。

为了更好地进行狩猎活动，采集-狩猎者会建造一些大型工程。驱赶线（Drive Line）是一种类似于篱笆的工程，主要用以猎捕驯鹿、水牛、大角羊、叉角羚等动物。有了这种工程，人们就能非常轻松地捕猎了。在大盆地（Great Basin）地区，这种被广泛使用的捕猎方法至少已有3 000年的历史。

在内华达州以北、莫诺莱克（Mono Lake）东部地区，有一种被称为“威士忌平地”（the Whisky Flat）的捕捉叉角羚的陷阱，无疑是一个极具研究价值的例子。[17]借助一个2.3千米长的“篱笆”，叉角羚被引入一个巨大的圆形围栏，而后被手持弓箭的狩猎者轻松捕获。这个围栏由5 000个左右的桧木桩组成。这意味着，许许多多的人得去砍树、挖木桩洞，以及收集和放置起支撑作用的石头。在大盆地地区，类似的驱赶线随处可见。平原印第安人修建了数千米长的驱赶线，将惊慌逃窜的水牛群引至悬崖边。[18]在加拿大的北极地区，因纽特人普遍使用驱赶线来猎捕驯鹿。他们的很多驱赶线都绵延数千米，由石块搭建而成，因而被因纽特人称作“石头人”。[19]在公元1 000年以前，生活在这一地区的多塞特人（Dorset）就修建了比较短的漏斗形驱赶线，或许是因为当时还没有弓箭，他们只能使用长矛更近距离地捕杀驯鹿。如图2-2所示。在阿尔伯塔（Alberta）和密歇根也发现了类似的石头驱赶线，而且在密歇根发现的驱赶线可以追溯到大约9 000年以前。[20]

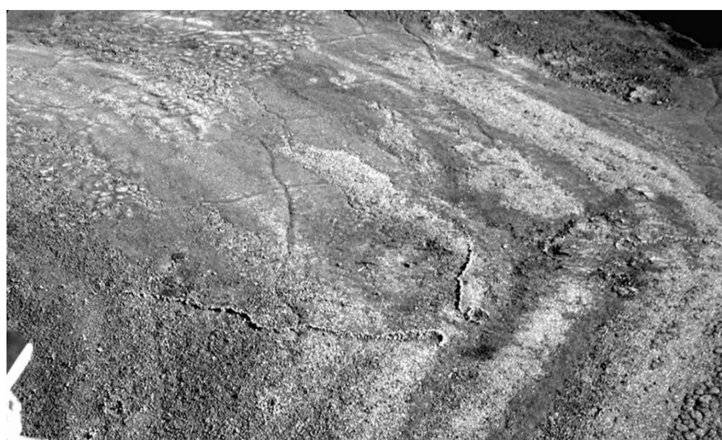


图**2-2**　加拿大努纳武特地区（**Nunavut**）维多利亚岛上的驱赶线遗址（**POD**）。两条驱赶线呈漏斗状，可在秋天将驯鹿群引入一个两边皆是陷阱的狭窄空间。陷阱里布满了长矛，这种设计是用来保护狩猎者的。紧邻陷阱的是实心墙体，长度在**17**到**40**米之间。驱赶线由一块块石头修筑而成。

From Brink, 2005; Friesen, 2013; image from Brink, 2005. Photo courtesy of Jack Brink.

沿河或沿海一带的觅食者通常都会修建堤坝，也就是水中的障碍物，用来捕鱼。大大小小的鱼堰遍布世界各地的大江大河。图2-3显示的是20世纪前半叶，胡帕（Hupa）人在加利福尼亚州北部修建的一个横跨米尔河（Mill Creek）的鲑鱼堰。这个堰是如何修建的？我们知之甚少。不过我们知道，大约在同一时期，尤罗克人（Yurok）在克拉马斯河（Klamath River）附近修建了一个类似的鱼堰。[21]人们将结实的木桩放置在河床上，而后筑起围栏，以拦截每年逆流而上的洄游鲑鱼。当鱼类强行通过鱼堰的狭窄通道时，人们就可以用渔网捕获大量的鱼。砍伐树木所需的人力在150人以上，修建鱼堰大约需要70人，然而令人惊讶的是，在使用了10天之后，整个鱼堰被完全拆除。大概是为了能让鲑鱼继续洄游，以维持与上游邻居的和平共处状态。

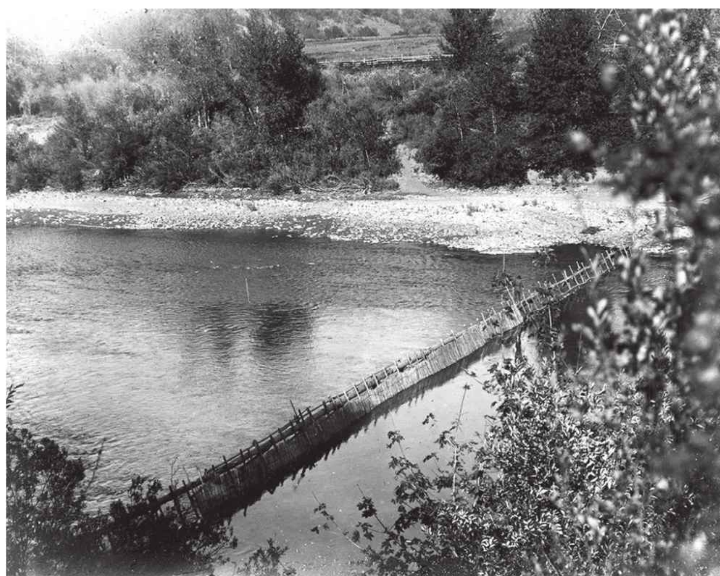


图**2-3**　胡帕人在加利福尼亚州北部修建的横跨米尔河的鲑鱼堰。

From Swezey & Heizer, 1977, 21. Copyright © Phoebe A. Hearst Museum of Anthropology and the Regents of the University of California, Photograph by P.E. goddard (catalog no. 15-3301).

群体间的冲突也会创造出重要的公共产品。很多现代战争就打破了参战双方之间的平衡，战败方会遭受大范围的掠夺和侵犯。在这样的战争中，成千上万的人面临着生命危险，而个体对战争的影响却微乎其微。英雄和懦夫同样都会体验到失败的悲伤和胜利的喜悦。这就意味着，在其他条件相同的情况下，利己主义者会尽可能地避开战争。当

然，在现代战争中，其他条件都是不一样的。为了激励士兵们投身战斗，现代国家还衍生出了极其复杂的制度。

狩猎-采集社会的战争，或者说群体间的冲突，是否也能创造出公共产品呢？这一问题的答案尚存在争议。争议的焦点是战争规模的大小。如果战争规模非常小，那么可能每个个体都能对战争结果产生实质性的影响。例如，负责在领地边界巡逻的黑猩猩通常有5至15只，它们一般不会轻易地发动攻击，除非在数量上远远超出对手。对于个体而言，如果发生冲突的成本非常低，那么随着领地面积的扩大，每个参与者从战争中获得的益处就会超出其付出的成本。

毫无疑问，狩猎-采集社会中也会出现群体间的冲突，而且死亡率通常很高。[22]然而，一些研究者认为，狩猎-采集者之间的冲突和黑猩猩之间的冲突是一样的，主要方式为伏击和突袭；冲突双方的人数都很少，且会尽量避免打响战争，除非其中一方在人数上占据绝对优势。[23]另一些研究者则认为，狩猎-采集社会也存在大规模的战争，双方都会派出大量士兵参战，最终导致大量伤亡。[24]目前的民族志资料尚不足以平息这场争议。到了20世纪，居住在阿齐族和哈扎族（Hadza）这类采集部落周围的农牧民都很有实力，因而对这些部族来说，发动战争是无利可图的。

在我看来，历史记载有力地证明了，采集社会中的人们有时候会发动大规模的战争，但参与者也会付出巨大的代价。最完整的数据来自澳大利亚。19世纪初，在欧洲人到来之前，生活在这块陆地上的人们一直处于采集社会时期。1803年，一个叫威廉·巴克利（William Buckly）的年轻人，因为接受了他人偷窃而来的衣物而受到了惩罚，被流放到澳大利亚某地。后来，他逃离了流放地，与澳大利亚土著瓦拉兰加人（Wallaranga）一起生活了35年。这段生活经历充满了大大小小的暴力冲突，比如谋杀、小规模袭击，以及全员参与的大规模战役。[25]那次战役的参战人数多达300人，最终双方都伤亡惨重。一份关于澳大利亚土著生活的专业论著证实了巴克利的描述：战争是非常普遍的，有时候规模巨大，死亡率极高，并会波及所有澳大利亚土著部落。[26]有关北美西部地区人们生活状况的历史描述也证明了冲突的普遍性。

据语言学相关资料显示，努米克语（Numic）的使用者，比如帕尤特人和肖肖尼人（Shoshone），在上一个千年里遍布整个大平原地区，不仅会和大平原周边的其他群体接触，还常常发起大规模的步兵战斗。[27]再后来，其中一个部落——科曼切（Comanche）获得了马匹，并对

相邻部落发动了大规模的骑兵突袭。最终，科曼切部落实现了霸权，领地范围从科罗拉多中部一直延伸到里奥格兰德（Rio Grande）。[28]考古学也证明，在白人移民进驻之前，大平原地区的战争是非常普遍的。[29]在西部地区，使用伊努皮克语系（Inupiaq）中不同语言的各个部落，也时常会发动大规模战争。[30]在加利福尼亚州的印第安部落中，尤其是在东南部沙漠地区以及沿海地区的部落中，战争是很常见的。[31]当然，小规模突袭也是很常见的。

此外，还有许多关于大规模冲突的报道，冲突双方的参与者均在数百名以上。在定居型的狩猎-采集部落中，比如新几内亚（New Cuinea）低地的采集西米的部落，和北美西北部沿海地区的捕捉大马哈鱼的部落，也常会发生数百人参与的战斗。人类学家A. L.克罗伯曾描述过加利福尼亚州西北部地区尤罗克人和胡帕人之间的一场战斗。在这场战斗中，约有100名尤罗克人对胡帕人发动了突袭。大概一年之后，为了复仇，100名左右的胡帕人向尤罗克人发动了突袭。[32]

我认为，关于狩猎-采集社会中的战争，历史材料证明了如下三个结论：

1. 在觅食部落中，暴力现象是普遍存在的。大部分的群体间冲突都以小规模突袭和伏击的形式发生。不过，在许多狩猎-采集社会中，也会发生双方参战者均达到数百名的战斗。

2. 在大规模的部落中，战斗规模往往受限于物资的补给，而非集体行动能力。在加拿大第一长河马更些河（Mackenzie River）三角洲西部的因纽特部落中，大规模的冲突时有发生；但在远东地区中部的因纽特人部落中，冲突却很少发生，而这些部落使用的都是因纽特语，且有许多共通的文化习俗。在西部，大规模的战斗更常发生，因为那里的人口密度更高。[33]

3. 大规模的战斗还会发生在使用不同语系语言的种族成员之间。如果游牧社会的觅食部落与现代社会的觅食部落规模相当，那么参与大规模战斗的成员必然会来自多个团队。据大部分资料显示，这些团队皆属于同一部落，而他们的对手则来自其他部落。但是，对于这种结论，需要慎重对待，因为一些欧洲人会从“民族-国家”战争的角度来进行观察，所以有可能会存在偏见。

非亲缘性合作的人类

我们已经了解到，人类的合作方式在其他脊椎动物身上是看不到的。任何一个人类社会的生存与发展都有赖于专业化分工和交换活动，以及公共产品的生产。例如，在过去的数千年间，游牧社会在这方面很可能与现代人类社会颇为相似。我们很少看到其他脊椎动物采用与人类合作类似的方式。不过，许多无脊椎动物，尤其是蚂蚁、蜜蜂、蜘蛛、白蚁等社会性动物，以及许多微生物，是会进行专业化分工和交换活动的，也会生产公共产品。在许多蚁群中，只有一少部分蚂蚁负责繁衍后代，其余大部分都属于工蚁或蚁兵，负责建造复杂的共同居住的巢穴。一些蚁群甚至还会培育真菌菌圃。诸如黏菌这类微生物也能进行大规模的合作。实际上，你可以将多细胞生物看作高度合作的细胞群体。在不同的机体组织中，细胞发挥着不同的作用，如吸收营养、安全防御和繁衍生息等。

在单细胞生物中，许多支系已经进化出了专业化分工与交换活动，栖息范围也有了很大扩展，这是生态学意义上惊人的成功。当来自不同分支的单细胞生物进化出专业化分工与交换活动时，多细胞有机体就出现了，其后代会占据非常大的细分生态位。这就好比群居性昆虫的食物占据了非常大的细分生态位，使昆虫既可以食肉又可以食草。这再次证明，无论在何种环境中，合作都能带来收益。此外，群居性昆虫会在许多领域进行合作。例如，蚂蚁会共同养育下一代；合作搭桥；共同守卫栖居地；合作管理交通，不同类型的成员各司其职。[34]

一些人认为，人类之所以具有合作性，是因为人类这一物种极具智慧。但现有证据表明，大脑尺寸并不是合作的必要条件。[35]狼群和非洲野狗也会合作猎捕并分享食物，但其大脑比人类的，甚至比猿类的大脑都小得多。其他懂得合作养育的哺乳类动物的大脑甚至更小。虽然群体规模与脑的大小呈正相关关系，但是许多群体规模较大的物种并不是特别具有合作性。诸如乌鸦、松鸦、喜鹊、渡鸦等鸦科动物的大脑尺寸与其合作水平并无关联。[36]其他群居昆虫，以及无脊椎动物以极其有限的认知能力在庞大的群体中实现复杂的合作。它们为了共同的利益而进行协作；不聪明但能够解决合作中的“搭便车”问题。

它们之所以能够做到这一点，是因为其合作建立在亲缘关系之上。许多进化生物学家都认为，可以将自然选择视为最大化的广义适应性（inclusive fitness）：用第一个个体与第二个个体之间的相关度，以数

量权重的形式，描述目标个体行为的变化之于第二个个体适应性的影响。相关度，也就是r值，描述了两个亲缘个体拥有某种相同基因的概率。直系兄弟姐妹，以及父母与子女的r值是1/2，祖父母和孙子女是1/4，旁系兄弟姐妹是1/8，以此类推。某种行为降低了一个人的适应性，用c表示降低的数值；同时增加了另一个人的适应性，用b表示增加的数值。当$r \times b - c > 0$时，行为的改变就是有益的。[37]这意味着，在以亲缘关系为基础的交往中，自然选择就像行为主体一样，对他人的适应性赋予了正值。[38]前文所提及的大部分社会性群体都是以亲缘关系为基础的，因而其合作行为的进化并不难理解。

在某些具有社会性的昆虫群体中，成员之间的相关度看起来比较低，但有足够的证据可以证明，这些支系在进化之初都是由关系非常密切的亲缘个体组成的。[39]有关繁殖的专业化分工一旦得以进化，工蚁就很难再繁衍自己的后代了，而对于那些无法繁衍后代的个体而言，合作的代价就会降低。最重要的一点是，专业化分工和公共产品生产在自然界中极为普遍；物种之所以能得以进化，是因为合作者之间存在密切的亲缘关系。

人类的合作是例外，可以在毫无关系的成员之间进行。就像米尔顿·弗里德曼所提到的铅笔，其制造过程需要无数人通力合作：加利福尼亚州北部的伐木工人、中国的采矿工人，以及华盛顿的工程师等。工程师操控着水电大坝，发出的电用于冶炼金属铝。公共产品的生产亦是如此。当代人类社会已经证明，人类具有这种非亲缘性合作的能力。

狩猎-采集社会中也存在同样的情况。人们通常会在史料中看到，狩猎-采集部落的合作一般出现在成员关系亲密的小团队内。然而，我的同事金·希尔及其合作者在两篇文章中明确指出，人种学对觅食部落所做的深入研究足以证明，上述描述是完全错误的。第一，在狩猎-采集部落中，合作团队的成员之间并不存在特别密切的关系。有一项研究调查了32个狩猎-采集部落，[40]结果表明，平均只有不到10%的团队，其成员之间具有密切的亲缘关系。阿齐族是一个被研究得颇为透彻的群体，就其60个团队而言，成员间的平均相关度指数为0.054。这就意味着，只有当收益为成本的20倍左右时，利他行为才会受到自然选择的青睐。第二，互利性的社会交往并不仅限于团队成员之间，而会延伸至同一语系其他群体的所有成员之间。[41]

如前文所述，狩猎-采集部落能够将不同团队的成员都动员起来，共同修建驱赶线、鱼堰和灌溉工程，而且还能召集数百名成员参加战斗。

关于阿齐和哈扎这两个当代觅食部落中的团队合作，有详尽的数据资料可以支持上述观点。北部的阿齐部落大约有500名成员，而哈扎部落大约有1 000多名成员。希尔及其合作者对阿齐部落的成员进行了走访，了解他们与同一语系其他群体的成员之间的交往情况。希尔展示了其他群体的成员的照片，并询问他们与这些人进行各种合作已有多长时间，比如分享食物和共同照料孩子。

耶鲁大学的人类学家布雷恩·伍德（Brain Wood）在哈扎部落中进行了类似的调查研究。[42]随后，他用这些数据测算出了每种合作行为的交互作用率（rates of interaction）。结果表明，平均来说，在阿齐部落和哈扎部落中，大部分成员会与同一语系其他群体的成员每两年合作一次。用外推法可以推测出，他们在一生中会和其他群体的成员合作很多次。当然，团队成员之间的合作要远多于“内外”合作。然而，团队成员具有非常强的流动性，个体经常会从一个团队转入另一个团队。大多数灵长类动物，包括各种类人猿在内，和平共处通常只存在于聚居群体内部，而不同群体成员之间的交往常常是带有敌意的。人类社会则不同。尽管人类的聚居群体和其他灵长类动物的有相似之处，但和平的合作性的互动常见于具有共同语言和习俗的群体之间。

在狩猎-采集社会的语系群体中，遗传学上的平均相关度是非常低的。在人口遗传学中，当群体规模非常大，迁移率非常低，且移民在定居前需要历经长途跋涉时，群体成员的平均相关度r可以通过如下公式来计算：

$$r \approx 1/(1+4Nm)$$

N是群体规模，m是每一代人的迁移率。[43]从这个公式中，我们很容易看出，就包括人类在内的大多数灵长类动物的社会群体而言，为何其平均相关度会如此低。对于像黑猩猩和狒狒这样的陆生物种而言，社会交往通常只发生在固定的群体中，且群体规模在100只左右。

某种性别的成员在性成熟期会向外迁移，同时在一个群体内大体上会存在两代人，因此每一代人的迁移率大概为25%，由此可得出$r \approx 0.01$。那么，在人类的狩猎-采集社会中，情况又是如何呢？社会交往通常发生在同一语系的群体内，群体规模通常在500人左右，迁移率约为0.05，最后得出$r \approx 0.01$。考虑到地方性移民、代际间的交叠、群体规模的变化，以及两性的影响等，预期校正值会更大一些。在最近的两份遗传学研究报告中，数值分别是$r \approx 0.03$[44]和$r \approx 0.07$[45]。这就意味着，如果个体从公共产品中所获得的收益，超出其支出成本的20倍，那

么公共产品的生产就会受到自然选择的青睐。

三四百万年以前的类人猿时代，“人类”的群体规模还非常小，来自不同群体的成员往往互存敌意，没有交换活动，也没有公共产品的生产。大约200万到100万年以前，现代人类开始出现，人们在同一语系群体内部进行合作和交换，以及公共产品的生产，群体规模在1 000人左右，即便是毫无亲缘关系的成员也能相互合作。除了人类的合作，亲缘关系还可以用来解释自然界中其他一切大规模的合作现象。由此，一个新问题摆到了我们面前：推动非亲缘性合作的人类心理机制是如何受到自然选择的青睐的？

互惠利他主义并不是解决之道

几年前，在一个论坛上，我与一位年轻的进化心理学家同台演讲。我在既定时间内粗略地介绍了一下人类是如何在“不相干”的群体间进行大规模合作的，并把这个主题作为一个进化难题提了出来。对此，那位进化心理学家说：“关于合作的进化，存在两种解释：亲缘选择和互惠利他主义。基于此，互惠利他主义必然可以解释人类的所有合作行为。”那些试图从进化的角度来理解人类行为的人们也持有类似观点，尽管没有直截了当地表达出来。

这位进化心理学家的观点听起来似乎是正确的，但实则不然，除非你把互惠利他主义看成是：个体基于其社会同伴的过往行为，于当前所做出的行为选择。然而，许多对人类进化感兴趣的人都将互惠利他主义看成是条件性合作（contigent cooperation）：如果你过去与别人有过合作，那现在我就跟你合作。不难看出，这种定义下的互惠利他主义只适用于小规模群体中的合作。大规模群体中的合作需要通过第三方来监管和执行相关的规范和制度。由此可见，这种带有强制性的规范与互惠利他主义是截然不同的。那么，从条件性合作的角度来看，为何互惠利他主义无法解释从合作伙伴那里观察到的合作行为的进化过程？对于这一问题的理解是非常重要的。

与互惠利他主义有关的进化理论最初出现在美国进化学家罗伯特·特里弗斯（Robert Trivers）的一篇缺少数据支撑的文章。[46]这篇文章的基本观点非常简单：我帮你，希望你将来也能帮我；如果在我需要帮助的时候，你没有提供帮助，那么我就再也不会帮你了。只有当长期合作所带来的收益超出短期欺骗性合作所带来的好处时，合作行为才会受到

自然选择的青睐。然而，只有当个体通过频繁的交往对合作伙伴有所认知，并清楚合作伙伴以往的行为时，合作才行得通。显然，互惠利他主义其实不具有丝毫的利他性，如果合作成功，那是因为合作者从中获得了长期的回报。因此，在这里，我将使用更为确切的术语：互惠。特里弗斯在其文章的最后提到，应该利用重复博弈的“囚徒困境”来建立互惠模型。

10年之后，美国博弈论专家罗伯特·阿克塞尔罗德（Robert Axelrod）和英国生物学家W. D.汉密尔顿（W. D. Hamilton）将特里弗斯的设想付诸实践。[47]他们采用了进化博弈论的研究取向，从人群中随机选取了许多参与者，经过配对后让参与者不断地进行互动。在每一轮互动中，参与者都可以选择合作或背叛。在生物学家所青睐的“囚徒困境”模型里，选择合作的人要付出适合度代价（fitness cost）$-c$，同时也会为对手创造一个适合度收益（fitness benefit）b，而选择背叛的人既不用付出任何代价，也不会为对手创造丝毫收益。在每一轮中，无论你的对手做出怎样的选择，背叛对你来说都是最好的选择。如果对方选择合作，则$b>b-c$；如果对方没有选择合作，则$0>-c$。如此一轮一轮地重复选择，而条件性行为使合作得以维持下去。

例如，假设有个体利用某种规则，也就是博弈论所说的“策略”，以牙还牙：在第一轮中，个体选择合作，而后根据对手在上一轮中的行为做出选择。当两个讲究策略的个体碰到一起时，他们都会选择合作，直到互动结束。当一个讲究策略的个体和一个经常选择背叛的个体相遇时，在第一轮结束之后，讲究策略的个体就会终止合作。如果讲究策略的个体普遍存在于人群之中，那么讲究策略者的主要交往对象也会是讲究策略者，而且会从持续的合作中获得长期的回报。尽管背叛者能够得到更大收益，但这种收益只是一时的，因为别人不会再与其合作了。只要合作时间足够长，长期收益一定会大于短期收益；与背叛相比，讲究策略的合作能获得更多收益。

阿克塞尔罗德与汉密尔顿的论文引发了大量有关重复博弈“囚徒困境”的研究，相关文献以“互惠的进化”为核心，从以下4个方面进行了论述：[48]

1. 当不相关的个体进行互动，且互动时间足够长时，在条件性合作普遍存在的群体中，个体可以避免由背叛带来的危害。这类策略被称为“进化稳定策略”，也就是进化博弈论中的ESS。时

间足够长，并非指非常长的时间。例如，如果$b \div c = 2$，那么个体平均只需进行两轮即可做出决策。尽管存在各种各样的合作策略，但如果互动能够持续进行，那么双方所采用的大部分合作策略就都具有进化上的稳定性。

2. 当不相关的个体进行互动时，非合作行为付出了代价，此时$c > 0$，那么互动也会在进化过程中稳定下来。

3. 如果进行互动的个体是有关联的，但关系不一定非常密切，那么在一个无条件背叛普遍存在的群体中，也会出现条件性合作。例如，如果$b \div c = 2$，双方进行8次互动，只要$r > 1/16$，条件性合作就会取代无条件背叛。其中，r指的是阿齐部落各团队内部的平均相关度。

4. 用一种合作性策略取代另一种合作性策略往往会更加困难。与取代背叛性策略相比，取代合作性策略通常会比较低效。[49]这就意味着，虽然人际间的相关性能够确保原始的互惠策略在缺乏合作的社会中持续存在，但互惠策略并不会因此而得到进一步的进化，从而变得更有效。

运算结果验证了罗伯特·特里弗斯根据语言学推理所得的结论：在社会性哺乳类动物和鸟类中，个体之间存在频繁的交往；个体能够识别每个群体成员，并且具有复杂的认知能力，因而互惠行为是非常普遍的。

制裁与互惠

用互惠来解释人类的合作行为存在两个大问题。第一个大问题是，该推论的前提是一种假设，也就是认为在具有复杂认知能力的社会性动物群体中，互惠行为普遍存在。但是大自然所展示出来的情况却正好相反。尽管还存在一些争议，但大多数行为生态学家皆认为，很少有典型实例可以证明，在其他动物群体中存在互惠行为，而灵长类动物只是特例；[50]即使能找到一些相关证据，那些证据也不具有足够的说服力。这就表明，该推论可能存在一些根本错误。在后文中，我们会再次探讨这一主题。第二个大问题更为突出，同时也是这个推论在解释人类合作时，尤其是在解释为何条件性合作只能为小规模群体的公共产品生产提供支撑时需要面临的问题。

要明白这一点，不妨把自己想象为澳大利亚的沙漠土著。你需要和

数百人共用几口水井，而每隔一段时间就会有相邻部落来抢占水井。每次遇到这种情况，你所在群体的所有成员便会联合起来，将侵略者驱逐出去。有时候，侵略者会还击，并导致战斗和伤亡。尽管会害怕，但为了保护水井，人人都参与了战斗，人人都认为这样的冒险是值得的。假如有成员在战斗之初就逃跑了，而其他人都投身于战斗，并将侵略者驱逐了出去，那么对于临阵逃脱者而言，战死的可能性大大降低了，而且事后依旧可以使用水井。由此可见，防御属于公共产品。

如果群体成员都是条件性合作者，都在采用“只要其他人投入战斗，我就投入战斗”的策略，那么只要有一个胆小鬼跑掉，其他人就都会临阵逃离。胆小鬼跑掉之后，形势会十分糟糕，因为所有人都不愿意保卫水井了。如果能明白这一点，那个胆小鬼或许会选择参加战斗。理论上，只有这样才能让合作稳定地进行下去。但事实上，这又可能会衍生出另外两个问题。当合作伙伴遇到不确定性或出现失误时，第一个问题就出现了。比如，一位勇士误以为撤退信号已发出，因而错误地撤退了；或者其他人一时间看不到勇士的身影，就误以为他逃跑了。这些状况经常会发生，而且每次发生都会引发撤退风波。第二个问题在于，策略如果很少出现，就不会受到自然选择的青睐。如果人人都是背叛者，合作就不会出现，除非群体规模非常小；即使群体成员都具有亲缘关系，合作也不可能达成。[51]那些允许背叛者存在的策略，可以帮助我们理解这个问题，但无法真正地解决这个问题。[52]

直接进行制裁，作用会更大一些。如果人们打算直接对胆小鬼进行制裁，而不是拒绝合作，那么制裁的形式可以有很多种，有一些代价高昂，有一些不必付出什么代价。合作者们可以对胆小鬼进行体罚，例如，澳大利亚的提维人（Tiwi）在惩罚违规者时就会把长矛刺在其腿上。[53]这种做法的代价颇高，不仅需要花费一定的时间，而且还有可能会遭到违规者的反击。合作者们还可以排挤违规者，不让其使用水井，也不让其享受领地上的其他任何群体资源，这是哈扎人采用的一种惩罚形式。[54]如果违规者想方设法地使用群体资源，人们就必须采取措施来抵制，这样一来，代价就变高了。最后，违规者可能会被剥夺某些权利，而这些权利通常都基于各种互惠行为。比如，当胆小鬼生病或受伤时，不会有人愿意照料他；打猎归来两手空空时，也没有人愿意把肉分给他。拒绝提供帮助，这种形式的制裁对那些潜在的助人者来说，在短期内是有益的，既让他们免受因遵循社会规范而遭遇的损失，又让他们维护了自己的声誉。[55]

这里出现了术语上的混淆，有必要加以澄清。有一些学者，尤其是进化生物学领域的学者认为，“惩罚”一词仅仅是指会导致短期代价的制裁，所以他们倾向于使用“负准互惠”（negative quasi-reciprocity）这样的标签来表示那种涉及克扣利益的间接制裁。[56]而另一些学者则用“惩罚”指代所有的针对背叛者的制裁，而且会清晰地表明这些惩罚是否需要付出代价。[57]我更喜欢后一种表述方式，因为它更符合“惩罚”这个词在日常生活中所指代的意义。在惩罚孩子的时候，我们往往会剥夺他们平时所享受的一些特权。这种做法有时候会让惩罚者受益，比如，当你惩罚自己十几岁的孩子，不让他出门的时候，你也就不用开车带他去商场了。不过，为了避免混淆，接下来只有当制裁者需要付出代价时，我才使用“惩罚”一词。

在公共产品的生产过程中，制裁之所以要比互惠更有效，主要原因有以下几个。第一，制裁的对象主要是背叛者，而失误或不确定性并不会激发一系列的背叛行为，这就意味着，在一个喧嚣的世界中，合作是可以稳定存在的。第二，少量制裁者的存在确保了合作的回报性。如果受到制裁所付出的代价大于进行合作所需要的成本，那么即便只存在为数不多的制裁者，其他人的合作性也会被激发出来。换句话说，即使群体中的制裁者数量非常少，人们的合作也能得到促进。[58]第三，制裁行为能起到威慑作用。人们因为害怕受到制裁而遵守规则，同时只有在因为失误而做出背叛行为时才会被制裁，而这种情况所需要付出的代价并不会太高。毕竟胡萝卜远远贵于大棒。如果使用奖励机制促进人们的合作，那么合作一旦完成，就必须支付报酬。第四，在互惠情况下，如果只有一个人被制裁，那么个体所付出的最高代价，仅限于终止合作以止损。如果制裁者实施制裁，那么制裁所造成的伤害会非常大。

上述这些论点同日常生活中一些众所周知的观念是一致的。在小范围的合作中，互惠发挥着重要的调控作用。如果你邀请我去你家吃晚饭，而我没有礼尚往来的话，你恐怕也不会再邀请我了。如果你参加了某个小社团，而一些社团成员不愿做出贡献，那么你以后可能也会跟他们一样。在大规模的合作中，起到调控作用的是直接制裁，而不是互惠。如果上万名工会成员集体罢工，人们就不会因为害怕遭到反击而停止搭便车的行为。因此，更为直接的制裁是有必要的。对于许多公共产品而言，情况同样如此。这些公共产品的提供需要由警察局、法庭之类的强制性机构进行监管和保护。

有不少学者对此提出了质疑。他们认为，在那些没有建立起强制性

机构的社会中，直接制裁，尤其是需要付出代价的直接制裁，在合作中起到了非常重要的作用。[59]那么，我们就来探讨一下这个话题，看一个被研究得非常透彻的案例：图尔卡纳（Turkana）部落的战斗。

图尔卡纳人为何而战

在采集-狩猎社会的战争中，固有的集体行动问题是如何被解决的？我们不得而知，因为当代的狩猎-采集部落通常不会参与大规模战斗。然而在很久以前，有一些部落，比如非洲的图尔卡纳部落则会参与战斗。这些战斗与狩猎-采集部落之间的战斗有着惊人的相似之处，可以帮助我们了解战斗的运作机制。

图尔卡纳人居住在干旱的东非热带大草原上，过着自给自足、居无定所的游牧生活，不断地从一个地方迁移到另一个地方，迁移的范围要比觅食部落的采集区大得多，而且还会“包容”许多“素不相识”的个体。这里没有集权，也没有强制性机构。那些发挥着领导作用的人，往往是凭借占卜能力、英勇表现，或者英明的迁移决策来获得声望的，尽管在群体行动中扮演着领袖角色，但并不具备公认的强制性权威。即便如此，图尔卡纳人还是会发动大规模的突袭，从同一语系的其他群体那里掠夺牲畜。

有关图尔卡纳人的信息，来自人类学家萨拉·马修的一项研究。我们一起策划了这项研究，但具体的研究工作全是她完成的。她对118名图尔卡纳男性进行了走访，这是一个非常具有代表性的样本。访谈主要涉及这些男性所参加的最近一次的突袭行动，最终统计出了88次不同类型的突袭行动，并获得了详细的资料。同时，她还搜集了他们对“勇士”行为和“懦夫”行为的各种反应。为了估算战斗的死亡率，她还对一部分男性和一部分女性的家庭结构进行了重构。[60]

通常，很多图尔卡纳人都会参与到突袭行动中，不过，一次被称为“绝密”（Stealth）的突袭行动却只有一部分男性参与。这次行动的目的是掠夺那些守护不严的家畜。另外，图尔卡纳人还会发动武装突袭，这种袭击行动不仅参与者众多，而且还会有战斗计划。如图2-4所示，每次武装袭击行动平均会有315名男性参与。

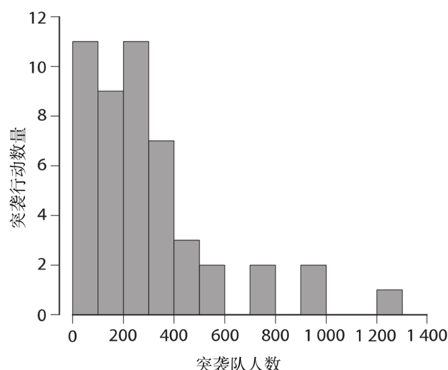


图**2-4**　参与武装突袭行动的图尔卡纳人人数柱状图。平均规模为**315**人，皆为男性。

Redrawn from Mathew & Boyd, 2011.

参战者必然会付出巨大代价。参与突袭行动的人不仅无暇顾及家人和牲畜，还有可能在战斗中受伤，甚至阵亡。通过对47次武装突袭行动进行分析，马修估算出突袭行动的死亡率是1.1%。马修通过走访搜集到的生活史资料得出，有20%的男性会死于战斗，其中54%死于袭击其他部落的行动，46%是为了保卫财产，死于由相邻部落发起的突袭行动。

这些大规模的突袭行动会产生集体利益，惠及每一位参战者，甚至每一位群体成员。如果突袭行动获胜，参与者们就会根据年龄，而非对战斗胜利所做的贡献来分配掠夺到的牲畜。有些突袭行动是由复仇之心所触发的，旨在报复之前袭击过他们的部落。这一类突袭行动所产生的震慑作用将惠及每一位群体成员。例如，如果一个部落在遭到袭击之后很快就发起反击，那么对方就会明白，这是在报复。参战者在出发时便很清楚，自己不可能因此而获得牲畜；即使两手空空地返回，这次袭击依然是成功的。大规模的突袭行动还有助于扩大游牧区域，从而获得旱季所必需的重要水源。在马修的走访样本中，规模最大的一次突袭行动发起得十分仓促，因为他们发现另一个部落的牧人迁移到了某处水源附近，而那里原本是图尔卡纳人的取水地。马修所掌握的信息表明，图尔卡纳人发起这次突袭行动是为了将入侵者驱逐出自己的领地。

图尔卡纳人的突袭行动具备了公共产品的各种基本要素。这些行动会带来集体利益，但同时参与者也要付出极大的代价。突袭行动的规模是庞大的，因而参战者个人对行动结果的影响会非常小，且付出的代价会超出所得收益。既然如此，图尔卡纳人为何还要参与战斗呢？当投机者拒绝参加突袭行动或者临阵逃脱时，为何整个系统并没有因此而崩溃呢？

事实上，投机取巧的行为的确会发生。每一位图尔卡纳男性都经常需要面对两难的选择，参加突袭行动或者退却。一般来说，突袭行动最初是由数位男性发起的，他们向群体成员传话，鼓励其他男性参加行动。男性成员有权拒绝参加突袭行动，但必须给出非常充分的理由。突袭队一旦组织起来，参战者们就会徒步数天直达敌人驻地。在行进途中，一些人会选择返回：在夜里偷偷溜走，或者告诉同伴自己生病了，或者说自己担心家里的牛群，抑或是说自己有种不祥的预感，觉得自己会死于突袭行动。调查发现，至少有43%的突袭行动都出现过参战者临阵逃脱的情况。

在战斗中，参战者有相当多的机会可以降低自身所面临的风险。突袭队在包围敌人驻地后便开始发动进攻，自此战斗正式打响了。一旦占据上风，年轻男性们就会迅速地将牧群驱赶到图尔卡纳人的领地内，而年龄较大的男性则继续进行“拖延”战，以阻断敌人的追赶。约有45%的武装突袭行动会出现激烈的战斗，据受访的参战者说，在战斗中，有的人会远远躲在别人后面，有的人不敢开枪，有的人一开始就临阵逃脱，有的人很快就撤退了。如图2-5所示。那些驱赶牧群的男性则会一直战斗下去，直到安全抵达自己部落的领地，然后等待其他参战者赶回来汇合。汇合之后，人们便开始瓜分战利品，然后回到各自的驻地。分配标准很明确，年长者比年轻人得到的份额多，同一年龄段的成员所得到的份额几乎相同。然而，在56%左右的武装突袭行动中，战利品的分配制度会被破坏，一部分参战者会夺走能夺走的任何牲畜。

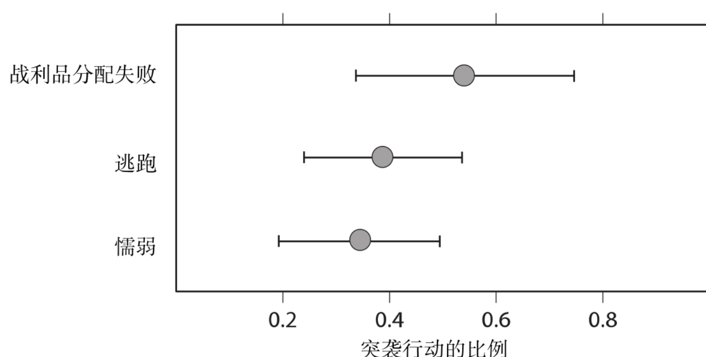


图2-5　出现各种投机行为的突袭行动的比例。

Redrawn from Mathew & Boyd, 2011.

然而，为了集体利益，图尔卡纳人年复一年地冒着生命危险参与着大规模的战斗。那么，他们何以能解决集体行动这样的难题呢？

原因并不是亲缘关系或互惠准则。突袭队成员相互之间并无关系，或者关系并不亲近。使用图尔卡纳语的人群被划分为20个不同的区域部落，这些区域部落又被称为“领地”，规模在2万人左右。这些部落属于父系氏族，同时还会按照年龄来分组，那些年龄相差5到6岁的男性成员会被划分到一组。尽管这些组相互之间时有联系，但频繁的日常交往主要还是发生在组的内部。平均而言，每次突袭行动的参战者都来自5个不同的年龄组、4个营地，以及3个部落。如图2-6所示。在被问及是否认识参加突袭行动的其他成员时，这些受访者的回答大多是：有些认识，有些不认识。

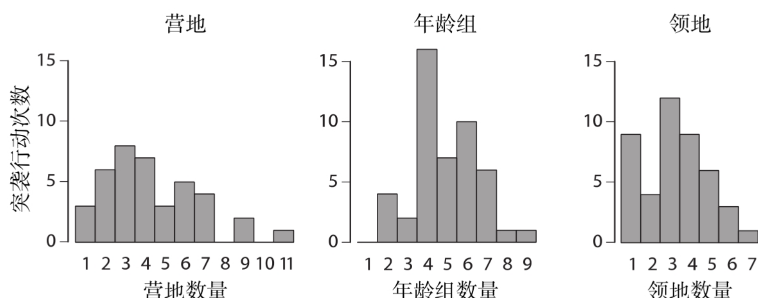


图2-6 突袭次数柱状图。纵坐标代表突袭行动的次数，横坐标分别代表营地、年龄组和领地的数目。这些突袭行动中的参战者都是从不同营地、年龄组和部落中选拔出来的。营地指的是一群人长期生活的地方，人群规模从几百人到上千人不等。年龄组是由那些接受割礼时间大致相同的男性组成，人数在**500**名左右。对于图尔卡纳部落的男性来说，年龄组是最重要的社会群体。部落成员的资格是通过继承和授权获得的，并附带有进入特定区域放牧的权力。

Redrawn from Mathew & Boyd, 2011.

制裁，帮助图尔卡纳人解决了战斗中的集体行动问题。约有47%的武装突袭行动出现了潜逃行为，而潜逃者都会受到制裁；在大约67%的武装突袭行动中，胆小鬼们都被记录在案，并受到了制裁，见图2-7。这些制裁是分等级的，从辱骂到罚款，再到体罚。体罚是非常严厉的制裁，违规者会被绑在树上，接受同龄人的鞭打。对违规者执行什么样的制裁，由群体共同决定：当地人，尤其是同龄人会对违规者的行为进行讨论。一旦达成共识，违规者所在年龄组的成员就要担负起执行惩罚的责任，即使这些人可能并没有参与突袭行动，或者并没有受到违规行为的影响。

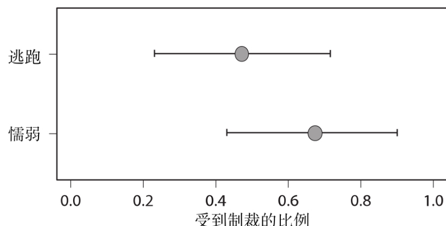


图**2-7** 不同类型的投机行为所占的比例。

Redrawn from Mathew & Boyd, 2011.

或许积极的刺激也会对持续性合作起到关键作用，比如，对那些在战斗中表现英勇的男性进行奖励。然而，关于这一点，很难得出一个明确的结论。勇士所获得的奖励通常都不是一次性的，很可能会终其一生不断地获得奖励。这些终身利益很难被测算。马修的研究无法解释，人们会基于何种因素来选择理想的伴侣或朋友，是某人在战斗中的行为，还是别的什么因素。然而，由直接制裁的比例可见，尽管间接制裁和奖励可能也很重要，但并非是个人价值得到提升的主要原因，如果它们具有足够的激励性，那人们就没有必要实施直接制裁了。

为何要实施惩罚

在小规模群体的人际交往中，关于“为何要实施惩罚”这一问题，有一个简单明了的答案。惩罚是为了制造激励机制，继而引导他人按照既定方式行事，以此提升惩罚者的福祉。

你决定惩罚一个盗贼，好让他再次偷窃之前会有所顾忌。在这个例子中，“你决定对盗贼进行惩罚”这件事和你未来的收益之间存在直接的因果联系。

在以大规模群体合作的形式来提升生产公共产品时，惩罚的实施似乎会遭遇一个难题。对于受罚者来说，惩罚是要付出代价的，但是对于每一个公共产品的受益者而言，惩罚又是有益的。惩罚可以激励人们对公共产品的生产做出贡献，但同时也会带来这样一个问题：为什么要实施惩罚？这就是“二阶搭便车”问题，一个相当令人困惑的问题。一些学者认为，这意味着惩罚不太可能出现；而另一些人则认为，惩罚是其他动机的副作用之一，或者是人类在小规模的群体生活过程中进化而得的某种心理机制所造成的结果之一。

我认为这一难题出现了错位。二阶搭便车难题的解决办法同一阶搭

便车难题的解决办法一样，也就是由第三方来实施规范性的惩罚。之所以要实施惩罚，是因为惩罚具有规范性。如果没有实施惩罚，就意味着要面临制裁。在有关惩罚的模型和实验中，人们需要面对的唯一选择是惩罚的程度，但在现实生活中，惩罚已植根于和道德规范有关的文化传播体系内。道德规范明确了哪些违规行为应该被惩罚，什么情况可以被缓期执行，相关证据是否充分，何种程度的惩罚是恰当的，谁有义务对惩罚进行监督，等等。假如，你负责惩罚某人，但你没有那样做，那么你就违反了规范；如果惩罚了一个无辜者，那么同样违反了规范；如果实施了过于严厉的、不合理的惩罚，那么还是违反了规范。如果违反了规范，便要面临制裁。正因如此，人们才有动力遵守规范。

在复杂的社会中，违法行为皆由国家强制性机构统一进行惩治。某些法官没有按照法律规定实施惩罚，或许是因为他们与被告之间存在某种私人关系，或许是因为被告有钱或有名。这些法官理应受到道德谴责，在制度健全的社会中，还应该锒铛入狱。马修所进行的研究表明，在图尔卡纳部落中，上述情况同样存在。[61]马修给图尔卡纳部落的一部分男性和女性看了一个十分简短的插画故事。故事讲述的是，在战斗中，参战者A目睹了参战者B的懦弱行为；在其中一幅插画中可以看到，A把这件事告诉了B的同龄人，而后大家达成一致意见，对B进行了惩罚。

这个故事的另一个版本是，目睹了B懦弱行为的A什么都没有做，而B也没有受到任何惩罚。针对不同的故事情境，马修会问他们，如何看待目睹了B懦弱行为的A？是否认为他做错了？他的行为是否会令你愤怒？是否有用？你是否会批评他？惩罚他？帮助他？你是否会跟他结婚？如图2-8所示，在得知B受到了惩罚时，图尔卡纳人对A的态度是非常积极的；在得知B没有受到任何惩罚时，他们对A的态度非常消极。马修还讲述了这个故事的其他几个版本，目睹B懦弱行为的A，对B实施了过于严厉，或者不公正的惩罚，对此，图尔卡纳人的反应就像得知B没有受到任何惩罚时那样强烈。

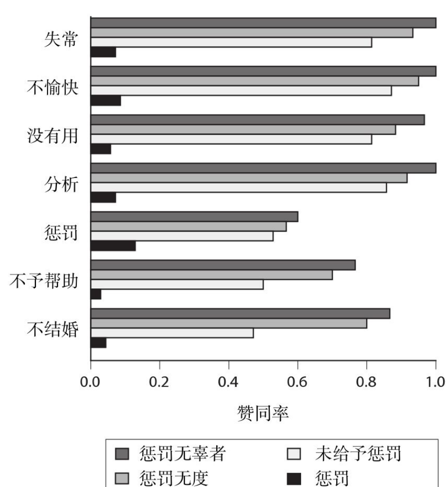


图2-8 图尔卡纳人在听到四个不同版本的故事时所做出的反应。该故事讲述的是一个人在突袭行动中目睹了另一个人的胆小鬼行为。在“有惩罚”的故事中，他在征求了胆小鬼所在团体意见之后，对胆小鬼进行了惩罚；在“无惩罚”的故事中，他什么也没做。在其他版本的故事中，他对胆小鬼实施了不公正的，或者过于严厉的惩罚。横坐标代表**60**名受访者做出不同反应的人数比例，纵坐标代表“不赞同”的反应。

Redrawn from Mathew, 2017.

有时候人们会认为，这种争论会导致无穷尽的递推。惩罚会带来二阶搭便车问题，而解决这个问题的办法是：对那些没有实施惩罚的人进行制裁。但是这又会导致三阶搭便车问题的出现：谁来制裁那些没有履行惩罚职责的人呢？而对这一问题的解决又会导致四阶搭便车问题的出现，依此类推，无穷无尽。对于许多人来说，这样的追溯似乎是不合理的。在我看来，关于这个争论其实有两个互补性的答案。

首先，行动者不需要追溯整个序列，只需要知道，在最后一个阶段是否违反了规范。如果某人的同龄伙伴没有履行惩罚职责，那么也没必要深究这一惩罚是否合理。更正式地说，在博弈论模型中，行为可以被设置在更小范围内的一系列状态变量中，这些变量概括了过去一段时间内的状态和行为信息，并会定期更新。[62]确定某人的行为是否真的违反了规范需要大量的背景资料，以及当时的特定情境等各方面信息。那么，追溯高阶违规行为是否比追溯这些信息更困难？这一点并不确定。

其次，当人们的关注点从二阶惩罚转移到三阶或四阶惩罚时，违规行为便会很少出现了。在平衡状态下，通常只有一小部分人不愿意合作，应该被惩罚；因此，只有小部分人需要承担起惩罚职责；相应地，只有极少数人不会履行惩罚职责。来看看接下来可能会发生的事情：随着一阶、二阶、三阶搭便车问题的相继出现，实施惩罚的预期代价也会

越来越低。这就意味着，针对n阶搭便车问题的自然选择力量会非常微弱，由此可见，即便是非常微弱的进化力量也能够起到稳定整个结构的作用。

稳定的社会规范

在我看来，人们执行规范的根本动机在于，认为“执行规范”本身即是一种行为标准。促使人们将规范内化并执行规范的心理机制受到了自然选择的青睐，因为人们在未能履行相应的职责，进而违反了所处社会的相应规范时，会遭到反对和制裁。我认为，一个违规者会受到严厉惩罚的社会会激发道德情操方面的基因进化，因此人类比其他灵长类动物更善于合作，更懂得信任，更愿意遵守和执行规范，也更少因犯规而付出代价。大量证据表明，人类具有其他灵长类动物所没有的道德情操，而且人类当下的合作在很大程度上是由道德情操所支撑的。关于道德情操的进化，在此不做赘述，不过我在其他地方有过深入的谈论。[63]

如果人们执行规范的主要动机是因为执行本身就是规范，那么随之而来的情况会是，规范的内容并不受执行机制的制约。同理，心理机制可以使无数规范稳定存在。比如，图尔卡纳人会发起或参与大规模的战斗；搭便车的人会受到同龄伙伴的惩罚；没有履行惩罚职责的人会受到制裁。很容易看出，只要受罚者付出了足够大的代价，图尔卡纳人的这一系列行为就能稳定地进化下去。但是，这种机制却没办法对图尔卡纳人的种种行为做出解释。

奖惩机制可以让许多行为稳定下来，包括非合作行为。相互进行制裁可能会使“可以偷邻居家的牛来供养家人”或“不可以偷邻居的牛来供养家人”之类的规范稳定下来。惩罚也能对那些“一定范围内受益”的规范起到稳定作用。“不要去偷氏族成员的牛，但勇士们可以去偷其他人的牛”，“不要偷自己部落成员的牛，但勇士们可以去偷其他部落成员的牛”，诸如此类均属于“对内有益”的规范，第一个规范使氏族受益，第二个规范则使部落受益。几乎所有的规范皆是如此。根据规范，必须把工作分配给最优秀的候选人，或者氏族成员；有时候规范要求必须娶自己兄弟的遗孀，但有时候又会判定娶兄弟遗孀的行为属于乱伦。此类事例，不胜枚举。

大量的行为通过进化稳定下来，由此可知，某些规范具有进化上的稳定性，然而这并不能给我们提供太多信息。如果要讲得更具体一点，

不妨参考一下建筑的结构。先从一堆砖块、泥灰和木料开始，利用这些材料可以修建出不同结构的砖式建筑，比如米兰圣安布罗焦（Sant' Ambrogio）教堂那样的罗马风格建筑、蒙蒂塞洛（Monticello）地区的乔治亚（Georgian）风格的宅邸、美国东部郊区随处可见的多层别墅，以及伦敦圣潘克拉斯（Saint Pancras）火车站那样的维多利亚式建筑。这些建筑在力学上都是相当稳定的。重力载荷和风力载荷，以及各种外力的相互作用共同塑造出了这些稳定的结构。

那么，知道圣潘克拉斯火车站是一个稳定的建筑，就可以解释其内在的结构吗？显然不能。虽然力学定律作用下的稳定性不可能支撑起著名儿童文学家瑟斯博士笔下的童话世界里的所有建筑，但足以支撑起现实世界中的各种建筑。但是稳定性本身并不能解释，为什么圣潘克拉斯车站会是这样的结构。规范，同样如此。知道哪些社会规范是稳定的，可以帮助我们了解当下的各种可能性；但未来会怎样，并不能由此而得出结论，或做出预测。

任何有关人类合作的真实描述都应该明确社会规范的产生过程，也就是明确和规范体系有关的文化进化过程。这是较为普遍的说法，用进化博弈论的术语来讲就是“均衡选择机制”。不同群体之间的文化竞争即属于此类机制。在下一节中，我将解释它是如何运作，以及如何影响我们在这个世界上所见到的规范的。不过我认为，其他的均衡选择机制也非常重要。我将会进一步思考这些问题。

文化的群体选择

达尔文在《物种起源》一书中提出，只要满足以下三个条件，自然选择就会促成累积性适应行为：

1. 生存竞争，以使部分个体得以繁衍。
2. 变异，以使某些类型的个体在繁衍速度上远超其他类型的个体。
3. 变异的可遗传性，以使子女与父母相似。

达尔文认为，上述三个条件同样适用于人类。他在《人类起源》（*The Descent of Man*）一书的第5章中提出，人类在社会性方面所具有的独一无二的特征会促使不同的群体选择不同的道德标准。虽然大多数

生物学家都很敬重达尔文，但许多人，或者说大多数人都反对以下观点：群体选择可以对合作行为的进化做出解释。达尔文思想的核心在于，只有当某些过程足以维持群体间的遗传变异时，群体选择才有效。在许多关注人类行为进化的学者看来，这种观点是缺乏说服力的。

彼得·里克森、约瑟夫·亨里奇、萨拉·马修，还有我，非常详尽地论证了达尔文在谈及文化进化时所持观点的正确性。[64]如前文所述，规范能起到激励作用，能使大量的行为稳定下来。通常，这些激励作用可以让相邻群体之间的规范差异得以稳定维持，即使群体成员和群体观念都在不断变化。这就意味着，随着时间的推移，群体间的规范差异会日趋稳定，哪怕群体产生分裂，其子群与母群也会保持高度的相似性。相邻群体之间往往存在着军事、经济等方面的竞争，而竞争是为了获得声望。群体间的竞争会促进某些规范的传播，但同时也会导致其他方面的损失。接下来我将提供相关证据，证明“文化群体选择”过程对人类社会的文化传播规范的形成起到了关键作用。一旦掌握了确切的证据就可以解释，达尔文的观点是如何与群体选择理论产生联系的，正如那些对社会进化感兴趣的进化生物学家们通常所理解的那样。

大量证据表明，文化传播能够让不同人类社会之间的规范差异稳定下来。19世纪末，人类学家在全球范围内对生活在殖民地的人们进行了研究。经过一个世纪艰苦且危险的工作，研究者们记录下了一系列令人眼花缭乱的社会体系。居住地仅仅相隔十几千米的不同群体可能会拥有完全不同的规范，而这些规范支配着人们生活的方方面面：跟谁结婚；把谁当作亲戚；在与他人发生冲突时必须支持谁；哪些行为是违法的；怎样惩罚违法者；等等。其中一些变异可能是环境差异所致，比如，与小规模的农耕部落相比，在游牧民族中，父系氏族制度更为常见。[65]

然而，正如我们在上一章中所说，大部分的变异源自共同的文化起源，而非环境差异。[66]住在沿海地带的萨利希人在行为上与住在遥远的内陆地区的萨利希人存在许多相似之处，而与近邻瓦卡希人的相似之处却很少，尽管其栖居环境与后者的大体相同。萨利希人之所以拥有能支配各种社会行为的一系列共同规范，是因为他们拥有共同的近代文化先祖，并从这些先祖身上继承了这些规范。文化研究所使用的系统发生法告诉我们，这一现象并不仅仅只发生在西部的印第安人身上。

群体中占主导地位的规范和价值观决定了群体将如何生存与发展：是否会被周边群体吞并；是否能吸引其他居民迁入。与那些可能会导致群体灭绝的规范相比，那些能够帮助群体生存下来的规范会被更广泛地

接受。同样地，那些能够促进扩张；降低被模仿的难度；吸引更多移民的规范会得到进一步增强，而那些不能促进发展的规范则不会得到增强。

人类学者理查德·索西斯（Richard Sosis）在19世纪至20世纪期间，对200个美国社会团体进行了研究，结果表明，规范会影响群体生存。[67]所有这些团体都不允许私有财产的存在，旨在支持公共产品的生产与使用。其中有88个团体具有宗教性质，其余团体都建立在某些强调集体生活的社会意识形态之上。如图2-9所示，具有宗教性质的团体要比建立在社会意识形态之上的团体拥有更高的存活率。最初，在这些团体中，仅有一半左右具有宗教性质，时至今日，那些存在了至少40年的团体几乎全都具有宗教性质。这就意味着，自然选择促进了宗教团体的发展，而非普通的社会团体。在成员资格方面，这些团体各具不同的规范。一些团体会对成员提出严苛的要求，比如要求成员不断祈祷，或者要求成员保持独身。索西斯发现，那些成员更多、要求更苛刻的宗教团体存在的时间要远长于那些对成员要求不高的社会团体。见图2-10。

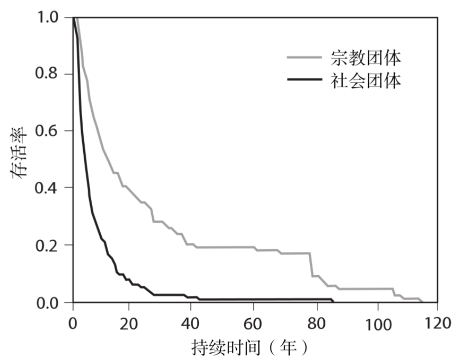


图**2-9**　社会团体的存活率与其存在时间的函数关系。黑线代表**112**个普遍社团，灰线代表宗教团体。

Redrawn from Sosis, 2000.

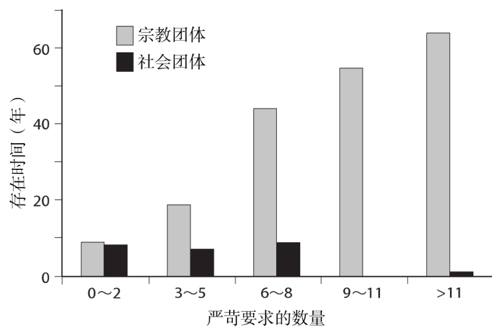


图2-10　社会团体存在的时间与对成员的严苛要求之间的数学关系。

Redrawn from Sosis, 2000.

这些数据足以说明，群体间的规范差异是如何得以维持的。不同的宗教团体拥有不同的规范。一些团体的规范非常宽松，而另一些则十分严苛。那么，在要求严苛的团体中，人们为什么要忍受那些要求？因为规范的执行是在群体监管下实现的，同时人们还害怕受到排斥和惩罚。由于规范的执行是由群体监管的，因此，不同的规范会对群体的生存产生不同的至关重要的影响。只有坚定的人才能加入要求繁多的团体，而在公共产品生产与消费所带来的诱惑下，这些团体得以幸存，且存在的时间比其他鱼龙混杂的团体要久得多。

这一结果与最近一项跨文化领域的实验的研究结果相同。约瑟夫·亨里奇及其同事对15个来自世界各地，拥有不同文化背景的群体进行了3项带有实验性质的“博弈游戏”。[68]在最简单的“独裁者”游戏中，每个参与者都有机会支配一笔钱：先给自己保留一定份额，剩下的需要分配给另一位匿名玩家。在这15个不同的群体中，某个类似于基督教的世界性宗教团体获得了23%的转让份额。

在19世纪的苏丹，努尔人（Nuer）在战胜丁卡人（Dinka）之后实现了扩张。这是一个关于“规范差异影响群体扩张”的典型实例。最初，努尔人和丁卡人各自为政，均建立起许多独立的部落。[70]在婚姻制度、彩礼规范以及政治制度等生存实践方面，努尔人和丁卡人鲜有共通之处，尤其是在政治制度方面，二者差异很大。受到地理因素的制约，丁卡人需要在雨季扎营并生活在一起，因而部落规模也受到了影响。努尔部落是依照男性亲缘关系组建起来的，规模是丁卡部落的3到4倍。当努尔部落与丁卡部落发生冲突时，努尔部落占据了优势，不仅因为其部落规模更大，还因为战争一般会发生在旱季，而在这个时候，努尔部落的营地规模要远大于丁卡部落的。冲突的结果是，13万左右的丁卡人被编入到了努尔人的体系中，最终被努尔人的规范同化。

这个例子说明，群体间竞争对文化群体选择是有要求的。群体之间必然会存在持久性的文化差异，而这些差异必然会对各个群体的竞争力产生影响。失败的群体必然要被胜利的群体替代，但胜利者无需将失败者统统杀掉，只需要进行分散或同化即可。

在小规模群体中，这类竞争是非常普遍的。来自新几内亚岛的数据便是最好的证明，早期的人类学家对当地传统社会的研究为后来的研究者提供了大量的样本。在人类学家进行研究时，这些传统社会尚未因欧

洲人的到来而发生重大转变。人类学家约瑟夫·索尔蒂斯（Joseph Soltis）借助许多早期人类学家所做的，与新几内亚岛传统社会有关的研究报告，整理出了与群体冲突有关的数据。[71]许多研究报告都对群体冲突的作用持肯定态度，约有一半的报告都提到了当地传统部落灭绝的案例。有5项研究提供了比较充分的信息，可以用来对相邻部落的灭绝率进行估算。见表2-1。

表2-1　位于新几内亚岛不同地区，拥有不同文化的5个群体的灭绝率

地域	群体数量	社会性灭绝数	经历了多少年	每 25 年灭绝率（%）
梅恩加（Mae Enga）	14	5	50	17.9
马林（Maring）	13	1	25	7.7
门迪（Mendi）	9	3	50	16.6
福尔 / 乌苏鲁发（Fore/Usurufa）	8 ～ 24	1	10	31.2 ～ 10.4
总计	26	4	40	9.6

From Soltis et al., 1995.

通常，一个部落在与相邻部落发生持续性的冲突之后，很可能会逐步衰落，直至被完全击败。有相当数量的成员会在确定本部落已不堪一击时，向其他部落中的朋友或亲戚寻求庇护。由此，他们的部落会转向社会性灭绝。这些数据为我们提供了，由文化群体选择所致的变化在概率上的上限。某种新的规范一旦出现，且那些不具有此种规范的群体渐渐走向了灭绝，那么只需经过20代人，或者说500年的时间，就会通过文化群体选择，从一个群体被传播到其他很多地方性群体中。这个过程看似非常缓慢，不过相较而言，以往那些庞大且复杂的全新世社会的发展历史可是经历了上万年的时间呢。

倾向于模仿成功者的心理机制，也有助于传播那些对群体有利的规范。人们通常会知道相邻群体的许多规范，比如，这里的人可以随意采摘果实，那里的人都拥有自己的果树。如果相邻群体普遍存在不同的规范，那么相较而言，某些系列性的规范可能更有助于人们获得成功。相当数量的证据表明，人们具有强烈的模仿成功者的倾向。[72]这就意味着，一些成功群体的信念可以通过各种渠道被传播到相邻的一些不太成功的群体中。[73]

最典型的例子莫过于基督教在罗马帝国的快速传播。在君士坦丁被征服前的280年间，基督徒的数量从一开始的少数，发展到后来的600万，再到3 000万，年增长速率达到3%至4%。按照社会学家罗德尼·斯塔克（Rodney Stark）的说法，许多罗马人之所以选择信奉基督教，是因为在早期的基督教团体中可以获得更高质量的生活。[74]其他宗教团体并无太多互助传统，不会特别关照贫穷之人或老弱病残者。当严重的流行性疾病暴发时，互助精神尤为重要，而罗马帝国在其后期，就因流行病的肆虐而惨遭重创。其他宗教徒拒绝帮助病人，甚至拒绝埋葬死者。最终，整个国家陷入了无政府状态。而当时的基督教团体则拥有强有力的互助规范，不仅能确保病患得到悉心照料，还能有效地降低死亡人数。

选择性移民也能促进某些有利于群体的规范的传播。在当今各个社会之间，移民潮频现。围绕这一主题[75]的大量文献皆支持如下两个基本结论：第一，人口往往会从那些不被看好的地区流向发展前景较好的地区。第二，经过几代人的时间，大部分移民都会融入迁入地的主流文化。据人种志记载，选择性移民现象古已有之。[76]同古代中国、古代罗马，以及古代印度等复杂社会有着密切联系的文化体系表明，选择性移民现象并非出自现代社会。古代帝国一般是通过军事行动来进行扩张的，但是像罗马帝国这种长期存在的帝国之所以会成功，部分原因在于它吸纳了大量的外来移民，而这些移民也接受了罗马帝国的社会规范体系。作为一个政治实体，罗马帝国最终还是消亡了，但其最具吸引力的制度，在被后来者改进后，以新的形式沿用至今。

明确地讲，文化的群体选择并不一定会带来进步。它能让那些有助于群体在竞争中获得成功的规范被广泛传播。约在1万年前，食物加工技术出现了，此后，那些有助于扩大群体合作规模的规范在世界各地逐步传开。这是否算是一种进步？见仁见智。社会规模的扩大为某一时期某一地区的人们带来了财富、长寿、安全保障和智慧启蒙。然而，在另一时期另一地区，文化的群体选择又可能会导致等级制度、剥削压迫、奴隶制度，甚至种族灭绝。在农业出现以前，人类至少拥有10万年的文化进化史，在这一过程中，那些可能让社会变得更小、更贫穷，但却更公平的规范反而更受青睐。而且，文化群体选择是一个相对缓慢的过程，有赖于长期存在的群体间竞争。当今世界，外部环境正在以疯狂的速度发生着变化，我们没有任何理由认为，现有的规范都能很好地适应当前环境。

为什么许多人拒绝接受文化的群体选择推论

希望你会认可这样一个观点：文化的群体选择推论是合理的。它不仅符合我们对人类社会史的认知，而且还拥有大量的例证。另外，在某种程度上，这些规范能对公司、大学以及其他复杂社会机构中的行为进行调控，以此在这些机构的进化过程中发挥重要作用。例如，不同企业在文化上的差异是不言而喻的。有些企业规范能够促进自由创新，有些企业规范则很保守和官僚。因此，一旦某些文化差异影响公司的生存和发展，并使公司成为其他公司的模仿对象，那么文化的群体选择就会对当下同类公司都产生影响。对于宗教团体、俱乐部、政体，以及复杂社会中其他存在文化差异的群体而言，同样如此。

这一切看起来合情合理，但实际上，文化的群体选择理论在人类进化学上是存在争议的。群体的基因选择理论是否有用？围绕这一问题，生物学家们进行了非常激烈的争论。事实上，群体的基因选择理论是非常有用的。该理论主要关注的是那些能够使群体受益，但同时要求个体付出代价的行为，其核心问题在于：群体选择能否战胜群体内部的选择？这一问题的答案主要取决于群体间遗传变异的比例。[77]增加群体间的遗传变异量会使群体选择具有决定性意义。如果群体间存在足够大的变异量，那么群体选择就会占据上风。这看起来似乎非常简单。

之所以会出现争论，是因为人们还可以用广义适应度或亲缘选择来解释这些群体选择模型。想要理解这场争论的基本特点，可以参阅史蒂芬·平克(5)的文章《群体选择的虚假诱惑》（*The False Allure of Group Selection*），以及相关网站上的评论。当大部分变异都来自群体间时，群体内部的相关度会增强，所以最大化的广义适应度会假设有益于群体的行为更可能得到进化。在解释同样的潜在过程时，这两种方法是一样的。如果计算准确，那么这两种方法得到的答案会相同。在我看来，这两种取向都是有用的，许多生物学家也都认同这一点。[78]但是仍有许多人认为，群体选择理论常令人感到困惑，而且还会让人产生误解，因为其本质上就是亲缘选择理论。[79]关注人类行为进化的大部分学者所了解的进化生物学方面的知识都来自持这种观点的人，而他们认为基因的群体选择压根儿就不重要。

文化群体选择则不同，它是具有不同社会分工的群体选择，在进化上很是稳定，并非是指那些使群体受益且要求个体付出代价的行为的进化。在这一点上，它与上述理论截然不同。正如我们所看到的那样，规

范可以使一系列行为稳定下来；同时，与移民或其他形式的融合相比，文化适应更加快速，因而具有适应性的文化学习过程可以促使规范将不同群体中的不同行为稳定下来。[80]这就意味着，在群体内部，不仅行为是由个体利益决定的，而且还存在大量的社会分工，同时每一种社会分工都与个体利益相一致。在这种情况下，群体间的竞争便会对社会分工起到决定性作用。

值得注意的是，群体间的变异仍然十分重要，不过，当前的群体变异主要是由适应性的文化学习过程，也就是文化的进化选择过程所致，并持续进行着。自然选择通常无法保持群体间变异的稳定性，或者说在这方面的作用微乎其微，不足以抵消移民所带来的影响。[81]在与社会进化有关的遗传模型中，群体间的变异主要靠随机的遗传漂变维持着。这一差别主要会导致两个后果：首先，在遗传模型中，群体间的变异量对群体规模非常敏感，在其他条件相同的情况下，随着群体规模的扩大，群体内部的相关性会逐步降低；而在文化模型中，群体间的变异量并不取决于群体规模，不过在大规模群体中，文化学习过程能够创造并维持大量的文化变异。[82]其次，当存在多重平衡，且适应过程的作用大于移民的影响时，群体可以在不同的频率上保持稳定。不过，就算可以从整体上了解群体的频率特征，也无法预测群体内部的相关性，而需要记录下每个群体在不同频率上的动态变化情况。

因此，当适应过程强大到可以维持群体间的变异时，要将广义适应度应用于文化变迁分析是比较困难的。据我所知，目前还没有人能够成功地完成这一类分析。事实上，进化生物学家斯图尔特·韦斯特（Stuart West）及其同事[83]曾为这类模型据理力争。在该类模型中，适应过程的作用很微弱，群体间的文化变异主要靠文化基因的漂变来维持，因此这类模型才可以用标准化的广义适应度来解释。唯一的问题是，这类模型与文化变异的实证模式并不一致。[84]相反，群体选择理论被用来模拟由群体间竞争所引起的平衡性选择已有30多年的历史了。[85]广义适应度和群体选择理论具有相同的效用，而我们理应选择最实用的方法。就当前情况看来，最实用的方法应该是群体选择理论。

若有机会重新开始，我不会使用“文化群体选择”这一术语，因为它会激起部分人的消极想法，在那些人看来，只有当个体与群体发生冲突时，群体选择才是重要的。我和人类学家彼得·里克森在20世纪80年代早期提出这一理论时，并没有预料到会出现这种观点。当时，我正在芝加哥大学生物学家迈克尔·韦德（Michael Wade）的实验室做博士后研

究。当我们提出，群体选择是一个很重要的过程时，周围人都表示赞同，韦德则是主要倡导者之一。

不仅如此，遗传学家塞沃尔·赖特（Sewall Wright）还提出，当不同的基因发生交互并生成多重平衡时，群体选择便会在基因进化过程中发挥至关重要的作用。他的观点着实令人振奋。赖特所设想的这一过程的数学模型，与文化的群体选择模型非常类似。[86]历经了30多年的争论之后，我现在终于明白了，或许当初应该使用人类学家萨拉·马修和马尔·泽弗曼（Marr Zefferman）所提出的术语：群体结构性文化选择（group-structured cultural selection），那才是明智的做法。[87]

文化的群体选择并非唯一的解释

我认为，群体选择理论所采用的机制是比较合理的，同时大量例证表明：群体选择促成了规范的形成。不过，在我看来，许多规范转变并不是群体间竞争的结果。决斗规范的转变便是其中一个典型例证。[88]直到19世纪初，在欧洲的上流社会中，和侵犯荣誉有关的规范要求受侵犯的男性要向挑衅者发起决斗挑战。于是，在1829年，时任英国首相的威灵顿公爵（Duke of Wellington）向温奇尔西伯爵（Earl of Winchilsea）发起了决斗挑战，因为后者在国会上讨论《天主教解放法案》（*Catholic Emancipation Bill*）时指责前者不诚实。这就好像，奥巴马觉得自己必须与参议员特德·克鲁兹（Ted Cruz）一决高下，以此捍卫自己的荣誉一样。威灵顿的枪法是出了名的差，在他失手后，温奇利西向空中开了枪。

此事件的25年后，决斗规范在英国彻底消失了。在过去的几十年间，美国社会中有关抽烟、婚前性行为、非婚生子女以及同性结婚的规范也发生了转变。所有这些转变看似都是由群体间竞争造成的，实则未必；同时，并没有任何部落消亡或扩张，也没有出现新的被人模仿的规范。

目前尚无理论可以很好地解释规范转变。历史学家对一些特例进行了看似可信的阐释。例如，政治哲学家夸梅·阿皮亚（Kwame Appiah）[89]认为，在英国的上流社会里，因为商业和手工制造业的利益变得愈加重要，所以决斗规范才会消失。

博弈论专家[90]所提出的模型类似于种群遗传学中的“峰值漂移”（peak shift）模型。[91]在该模型中，统计结果之所以会波动，是由

于有限的人口规模所造成频率的特征会随机游走。处于某一种平衡状态的种群人口会随机游走至另一种平衡状态，一段时间之后又会返回原来的状态。

我并不认为这些模型是符合实际情况的。首先，在大规模群体中，等待规范转变的时间会相当漫长，尤其是在“一旦偏离现有规范即会付出巨大代价”的情况下。其次，规范的转变是可逆的。朝某个方向转变要比往相反方向转变更容易，但每个种群内部最终都会反复发生多次转变。从模型中的变化来看，出现上述情况的可能性很小，而且群体越大，错误率越低；同时在规范转变发生之前，等待的时长也会越长。

人类的小规模合作

到目前为止，我一直专注于大规模群体的合作。然而，人类同时也是小规模群体中的特殊合作者。劳动分工和交换活动会发生在每一个人类社会中。如前文所述，在觅食社会的经济活动中，交换活动发挥着至关重要的作用，不仅降低了人类的死亡率，还使人类的生育周期比其他灵长类动物短了很多。[92]在其他动物群体中也存在许多小规模合作，不过这些发生在非亲缘个体之间的合作似乎都只是一些低成本行为，比如相互梳理毛发的行为。[93]

由此可见，互惠并不足以解释人类的小规模合作。许多其他的脊椎动物也生活在稳定的社会群体中，也能记住同类过往的行为表现，并据此来调整自身行为。如果仅靠互惠就能使人类的交换活动稳定下来，那么我们理应也能在那些动物身上发现更多的交换行为才对。然而，事实并非如此。

另外，在其他哺乳类动物的群体中，小规模合作也仅限于低成本行为，这一事实同样表明，互惠并没有发挥决定性作用。互惠模型的关键之处并非是成本，而是收益与成本之间的比率。这便意味着，互惠可以解释食物分享之类的行为，这种行为给受益者带来了更大的收益，而代价由施予者承担。例如，当受益者因受伤而无法自理时，或是因分娩而异常虚弱时，合作的收益就会远超所付出的成本。这些皆属于日常事件，相互帮助是很寻常的现象。作为独特的哺乳类动物，人类会经常分享食物，而每当此时，人们的收益都是非常高的。[94]

接下来我将证明，规范使人类比其他物种拥有更多的小规模合作，因为规范能够解决监督和执行的相关问题，而正是这些问题限制了其他

物种在互惠行为方面的进化。这一观点最初是由萨拉·马修提出的，而后我们又与人类学家马斯吉斯·范维伦（Matthijs van Veelen）合作，对这一观点进行了更加充分的探究。[95]

规范调控人类的小规模合作

由进化论学家提出的和社会生活有关的论述，往往带有自由主义色彩。他们将“社会”定义为：在自私自利的个体和任人唯亲的家族之间形成的一种双边交易网络。制裁和道德就此出现，因为行动者试图以增加自身利益为目的来改变他人的行为。一些研究者承认，规范对于大规模合作而言是不可或缺的，但其中的大多数人仍然认为，对于小规模合作来说，规范就不起作用了。

然而，这样的看法并不符合实际情况。人类社会中的小规模合作通常由第三方制裁所实施的共同规范来调控。个体不能随意进行交易，事务的处理必须遵循既定规范。而且，这些规范还会对与适应度密切相关的行为产生影响，比如结婚和生儿育女等。大多数社会都遵循着“婚姻”这种制度化的两性结合方式，而且婚姻定然会与既定的权利、义务相联系。我们经常会看到这样一些规范：禁止来自相同社会单元的成员结婚，只允许来自不同社会单元的成员结婚。这些规范往往会夸大配偶之间的潜在利益。

澳大利亚西北部沙漠中的瓦尔比利人（Walbiri）被划分为八个部落，部落成员只能与其他七个部落中的一位成员结婚。[96]规范还会明确指出，一个人可以与多少人结婚。在实行一夫一妻制的社会中，一位男性是不能同时和两位女性结婚的，即便他们认为这样做更符合其长远利益。规范还会对结婚所涉及的财产转移问题做出调控，因此在一些社会中，男方需要向女方家庭支付彩礼，而在另外一些社会中，女方需要向男方家庭支付彩礼。婚后的住处也会受到规范的限制。在“从夫居”的文化背景下，男方是不能与女方及其家人住在一起的，否则就会很丢脸。

规范还会对生儿育女的方式进行调控。在国家级别的社会中，规范会限制教育和训练孩子的方式。这方面的规范存在非常大的文化差异，例如，父母鞭打孩子的哪个身体部位属于违法，以及在什么情况下父母不愿鞭打孩子即是失职，等等。在过去的许多社会中，送女孩去学校读书是有违规范的，这种情况直到近代才有所改变。这种规范导致相当多的女性被排斥在劳动力市场之外。亲代投资也会受到继承制度的约束，

因为继承制度明确规定了应该如何在多个子女间进行财产分配，比如只分给儿子或女儿；分多少给长子，分多少给幼子；或是平均分配给每个孩子。在遵循平均分配制度的社会中，长嗣继承制是有违规范的；在遵循平等继承权的社会中，剥夺女儿的继承权则有违规范。

朋友之间的交往属于直接互惠的典型例证。尽管如此，对于图尔卡纳人而言，来自群体的支持和反对，为朋友间的持续合作奠定了重要基础。[97]比如，一名妇女如果拒绝向一名口渴的男性提供饮水，就会受到指责；但是，如果她听说，这名男性曾与朋友一起潜入敌人领地偷窃牲畜时放弃了受伤的朋友，那么她大可拒绝向其提供饮水，且不必担心会受到周围人的指责。同样，当一名牧民长途跋涉，四处寻找自己丢失的牲畜时，他或许会得到某个图尔卡纳家庭的款待；但是，如果这个人之前偷窃过邻居家的骆驼，那么就会被拒绝入住。邻里之间的关系除了会受到交往历史的影响，还会受到社区的管控。

当个体之间发生矛盾时，规范也会发挥重要作用。某个图尔卡纳男性在羊被偷了之后会把此事报告给长者，或告诉家族成员们。然后，他们会叫来嫌疑人，并商议接下来该作何处理。如果嫌疑人确实犯了错，他们就会责令其向丢羊之人做出赔偿。

请不要误会我在这里所说的话。人类并不是只知道按规范行事的机器人，而是需要维护个人利益的行动者，进行交换活动是为了增进自身利益。然而，规范会影响交换活动的类型。首先，人们会对违规所需付出的代价做出考量。在非常看重“性忠贞”的社会中，未婚双方通常不会忽略可能会遭遇的严重后果而接受所谓的“开放式婚姻”。在反对一夫多妻制的社会中，一个贫穷的女子不会忽略自己可能会遭受的指责而选择成为一个富豪的妾室。19世纪伊利诺伊州（Illinois）的人们对摩门教多元婚姻的态度就是最好的证明。

其次，我认为更重要的是，人们会将所处社会的规范内化，这样一来，倾向就会受到影响，行为选择也会受到影响。比如，在一些社会中体罚孩子是违法的，大部分人都认为体罚孩子会导致非常糟糕的后果，如果因一时情绪失控而动手打了孩子，那么事后也会感到非常羞愧；而在另外一些社会中，人们会将一些和处罚孩子有关的观念内化，反而会在狠不下心对孩子做出处罚时感到羞愧。[98]

第三方的监督和制裁可以稳定互惠关系

到目前为止，相关争论尚都处于经验主义的范畴内。人类展现出了多种形式的小规模合作，比如专业化分工和交换活动，而这些合作在其他物种中是看不到的。在人类的小规模合作中，共同规范发挥着至关重要的作用。此外，规范和第三方制裁之所以能够使有利于小规模互惠行为的条件在范围上得到扩展，还因为其他一些非常好的理论缘由，接下来我们将讨论其中几种。

当感知错误普遍存在时，来自第三方的监督和执行有助于稳定互惠关系。

互惠进化的标准模型假设，互惠关系在大自然中普遍存在。在这些模型中，类似“以牙还牙”的视条件而定的策略能够使无关个体之间的合作稳定地持续下去。

$$(1-1/T)\,b-c>0$$

这里的T指的是合作者之间的平均交往次数，b和c分别指的是受益者在合作中的适应度增量和施予者在合作中的适应度增量。[99]需要注意的是，该不等式的左侧部分表示的是广义适应度。在某种行为中，受益者的适应度增量是用（$1-1/T$）来衡量的，而非相关度。要想明白这到底意味着什么，需要先假设个体之间的平均交往次数为2次。然后，依照互惠进化的标准模型，只要收益超出成本的2倍，那么视条件而定的合作就会稳定地持续下去。氏族成员之间的合作也有赖于这样的收益成本比。如果个体之间的平均交往次数为10次，而收益仅比成本高出10%，那么依照此标准化理论，互惠行为理应非常普遍，并会促使那些与亲缘合作相比，收益成本比更低的合作出现。显然，这样的推断是大错特错的。相关度理论可以解释自然界中很多的高成本合作，然而互惠理论能解释的合作则非常少。

标准化理论有许多假设，而这些假设或许可以解释理论与实际情况之间的差异。我认为最主要的原因在于，标准化理论低估了误差的影响。同时，大部分此类模型都包含合作和背叛这两个选项。

在现实生活中，不同的成本和收益之间存在许多可能的交互项。合作的一方可能会考虑另一方的行为是否算得上是真正的背叛，比如，女孩决定去滑一天冰，而不是去机场见男孩；或者会考虑另一方的行为是否只是一种失误，比如，女孩原打算去见男孩，但是车子陷到了雪地里，一时无法解决。这就意味着，是否存在合作或背叛，对于处于互惠关系中的不同个体而言，很容易产生分歧。此类误解被称为“感知错

误”。如果这类错误很常见的话，那么互惠行为将会受到严重影响。

解决之道是采用一种特殊的策略：“赢定输移”（win-stay-lose-shift）。该策略可以让人们从感知错误中清醒过来。[100]然而，和其他许多互惠策略有所不同，这个策略并不会频繁出现，因为它很容易被背叛者利用。即使是在实验中，当个体出现感知错误并开始遵循这类策略行事时，这类策略也不会被表现出来。[101]

因为降低了错误率，加大了对背叛行为的惩罚力度，所以感知错误变得很常见。在这种情况下，第三方的监督和惩罚能够起到稳定互惠关系的作用。假设群体成员会对他人行为进行监督，并评判其行为是否符合规范，然后对观察结果进行汇总，并最终达成共识。在群体共识的基础上，人们会决定某个个体是否应该得到帮助。对于未参与监管的个体，人们会取消帮助，以此作为惩罚，这样便可以解决二阶搭便车这个难题了。[102]这样的第三方监督和惩罚可以让互惠行为在更大范围内持续存在，其原因有二：第一，增加监督者的人数可以增加违规者的曝光率，并降低忠诚者被误认作违规者的可能性；第二，增加实施惩罚者的人数可以使违规行为逐渐失去吸引力。建模研究表明，这些因素的影响力是非常大的。[103]

第三方裁定也可以增加互惠关系的稳定性，因为它为个体提供了一种机制，使人们在遭遇误解后还能重新建立起合作关系。比如，我和朋友约好在机场碰头，但她因为被困在雪地里而失约了。我的朋友会认为自己选择的是合作，但我可能会认为她选择的是背叛。于是，当她日后需要帮助时，我会拒绝。我的朋友会认为自己毫无过错，继而又对我进行了一系列的报复。现在假设，第三方对最初的争议进行了裁定，并对我的朋友是否真的选择了背叛做出判断。这样一来，个体便会得到一个非常清晰的公众信号，并可以依据这一信号对后续行为进行调整，从而重新建立起长期的互惠合作关系。

建模研究表明，这种调适可以极大地扩展互惠进化的条件范围，因为它能够使交往的个体在观念上保持一致，尽管这种调适过程本身也非常容易出错。[104]如果行为规范需要群体成员共同遵守，那么最容易调适的便是“一对一”的交换活动，因为它不会受到特定合作关系的制约。若非如此，在裁定纠纷时，第三方就必须对“是否偏离了特定规则”进行评估。

在地方性生态环境中规范可以确定如何进行合作。

人类的生活环境是非常复杂的，拥有多样化的生态系统、栖居模式以及社会制度。这就意味着，管理互惠行为的实际策略是需要通过学习来掌握的。例如，在狩猎-采集社会中，人们经常会分配肉食，但这并不只是一个有关“如何分配同类资源”的问题。对于不同的人，比如老人和年轻人，男人和女人来说，动物身上不同部位的肉具有不同的价值，有的脂肪多一些，有的蛋白质多一些。互惠策略很难借助通用的学习机制来掌握。在一些经过简化处理的实验室环境中，研究者们试图教会其他一些动物依照互惠策略行事，但鲜少成功。[105]与其他动物相比，人类被认为具有更强的学习能力，但即便如此，当不存在任何误差，且合作建立在进化稳定策略上时，人们仍然很难协调好各种合作策略。[106]因为现实世界中存在很多偶然性，因此对于某个特定的群体来说，要将注意力集中在某个可持续存在的互惠策略上则更加困难。

与互惠策略有关的文化进化规范有助于解决这一难题。在觅食者分配肉食的行为中，互惠策略确实发挥了作用，[107]但这类行为往往会受到具体的文化规范的调控。比如，这些规范会明确指出，哪些个体会分配到动物的哪些部位。例如，在刚果盆地的姆班吉拉（Mbendjele）部落中，相关规范明确指出，狩猎者可以分得动物心脏，其他男性可以分得动物肾脏，猎狗可以分得动物肺部，剩下的部分平均分配给其他人。这样一来，人人都能沾上狩猎者的好运气。[108]这种文化进化规范为人们提供了合作的基本框架，既能指导分配策略，又能限制冲突范围。

一位人种志学家曾给我讲过一个故事，与这种情况不谋而合。他曾经想要将一小部分烟草分发给研究对象，然而令他甚为苦恼的是，他发现那些人有太多的抱怨和欺骗。于是，他向一位当地人请教，自己是否可以分发这些烟草。那人阻止了他，并以一种急切的口气说：“千万别这么做，否则我们会相互残杀的。”这些觅食者每天都会分配肉食，并深知应该如何分配肉类，因为他们拥有经由文化进化而来的规范，明白在何种情况下哪些人应该得到哪些东西。然而，他们在分配某种有价值又很新奇的事物时会变得十分疯狂。

规范可能会强制施行配对交换，但有利于群体而非交换活动双方。

相较于子女自己，父母对子女之间的合作更感兴趣。不难想象，一对身处逆境的姐妹应该会相互帮助，姐妹中的任何一个在受伤时会得到另一个的照料。但广义适应度表明，只有当受益者的边际效益高出提供帮助者所付出成本的两倍时，她们才会相互帮助。然而，父母对子女之间互助性的期望值远高于此。在他们看来，如果可能的话，只要边际效

益大于边际成本，子女们就应该相互帮助。利益冲突会导致父母和子女发生激烈的斗争，而且结果如何很难预料。

然而，在氏族成员之间，如果互相帮助能成为群体的共同信念，那么这种规范便能发挥更好的作用，因为从总体上来看，一个互帮互助的群体可以让每个人都会过得更好一些，只要那些帮助是让合作双方受益的。在群体内部，人们皆会遵守规范，因为违规者会受到制裁。在正常情况下，有利于群体生存发展的规范会得以传播，从而让群体更具竞争力。

规范心理是如何得以进化的

我和萨拉·马修认为，无论是在大规模合作中，还是在小规模合作中，人们都会受到经由文化进化而来的共同规范的调控，而这些规范是通过第三方的监督和制裁来执行的。许多进化论学家则认为，小规模合作的进化是基因相关度和直接互惠行为所致。因此，唯一重要的事情就是，这对个体及其家庭是否会有影响。我的朋友乔和莫签订了一个对我似乎不太公平的协议。我为什么会这么在意这件事呢？关键是，我为什么要对此采取一些行动呢？我的邻居艾丽斯打了她的孩子，这原本不关我的事，但是生而为人，确实又会在意这些事。在社会中，自私自利并不是人类采取某种行为的唯一原因。我们会判断他人是否违反了规范，并会据此做出行为反应。当然，自私自利确实是重要原因之一，但与此同时，社会规范也在塑造和影响着人类社会生活的方方面面。为何会这样？

在我们看来，规范以及对违规行为的惩罚，为人类的合作养育提供了新形式。根据这一假设，在制作复杂的工具，以及加工食物以去除毒素等其他适应性的推动下，那些对累积性的文化进化有利的能力皆得到了进化。规范衍生出了解决各种问题的最佳方法，并以文化的形式进行传播，比如分配肉类的最佳方法。这样一来，在规模相对较小的群体中，与第三方监督和惩罚有关的收益推动了规范心理的进化。在人类早期社会中，这种规范心理促成了更加多样化的小规模合作，并且有助于弱关联群体从社会交换活动中获益。

道德心理因此而得到了进化，进而又通过文化群体选择促成了大规模合作的进化。文化的群体选择模型通常会假设，个体通过模仿成功者的行为或主流行为获得复杂的规范性行为。这也就是说，在学习其他地

方性适应行为时所采用的学习机制，同样可以用来学习当地的道德规范。不过，如果缺乏某些先天因素作为支撑，复杂的道德规范可能不太容易被掌握。在某些情况下，利益会趋于一致或产生冲突，因而我们在识别不同状况的相似之处时，对规范的思考大都十分抽象。于是，道德机制变得更加复杂了。在新环境中，人们通常能够识别出常见的因果关系，也就是利益冲突，并将其同规范联系起来。然而这样的方法无法依靠通用的文化学习机制来维持。小规模合作如果能够通过规范和第三方制裁得以稳定存在，便能为先天的道德直觉的进化提供支撑，从而使文化群体选择通过这种道德直觉轻松地促成大规模群体合作。

通过这种假设所得出的结论与我们所观察到的，各种规范对小规模合作和大规模合作的调控作用是一致的。在小规模合作中，对规范的遵守和执行取决于家庭成员所提供的线索，以及交往历史所提示的线索。我是否应该帮助乔？他是我的亲戚吗？我以前需要帮助的时候他是否帮过我？要实现对规范的遵守，就需要人们进行协商、深思熟虑，并达成共识。在大规模合作中，与环境变异有关的文化群体选择对规范的形成起到了更加重要的作用，对规范的遵守取决于和群体成员身份有关的线索。比如，我是否应该帮助琼？她是否和我属于同一族群？她是美国人吗？在这种情况下，即使缺乏协商和共识，规范也可以得以遵循。

文化让人类成了与众不同的生物

让我们将人类进化的录像带回放到400万年前。除了行走方式较为独特之外，人类的祖先与其他猿类拥有颇多相似之处。他们生活在撒哈拉沙漠以南的非洲大部分地区，群居方式也与其他猿类相似，还没有出现大规模合作，专业化分工和交换活动也很少见。

现在，我们快速地来到大约10 000年前的全新世之初。现代人类已经成为自然界中的佼佼者。与其他哺乳类动物相比，人类的足迹已遍布全球，能够适应各种各样的环境，能够利用各种各样的食物，并形成了各种各样的群体，而且比其他物种更具合作性。就小处而言，人类依靠专业化分工和交换活动养育出许多成长缓慢、生存成本很高的后代；就大的方面来讲，公共产品生产让更多人受益。当下，我们正在探究这一社会巨变的成因和途径。

我的基本观点是，无论是生态学意义上的成功，还是合作能力上的突飞猛进，都离不开文化适应性的进化这一最基本的因素。它虽然不是

唯一的因素，却是至关重要的因素。人类之所以能遍布全球，是因为具有经由文化进化而得的环境适应能力。在澳大利亚中部地区，制作大柄苹糕点和编织渔网是最基本的生存技能；而在北极地区，制作驯鹿皮衣和海豹油灯是同样重要的能力。[109]没有其他哪种生物具有如此多样化的环境适应能力。或许人类的确比其他动物聪明，但应对复杂环境的能力远比个体的创造能力更重要。这是渐进的文化累积过程所带来的结果之一。这个结果改变了一切。文化知识的累积有赖于人类大脑中固有的权衡能力。

在人们出于自身动机而采纳他人观念时，文化累积所起的作用最为明显，即便那些观念和自身经验是冲突的。基于此，可以说人类在一定程度上是文化的产物。人们之所以会相信自己愿意相信的东西，做自己愿意做的事情，一定程度上是因为周围人都相信这些东西，都在做这些事情。正因如此，社会规范才得以影响行为。只有当个体行为受到由第三方强制执行的规范的影响时，无关个体才有可能进行大规模合作。要想组织起数百人共同修建一道具有5 000个篱笆桩的驱赶线，必须有所激励，要让这些人觉得自己的付出是值得的。修建鱼堰和灌溉工程亦复如是，而最重要的莫过于同相邻部落的战斗。人们需要共同遵守一系列规范，比如什么样的行为是有必要的，由谁执行制裁，如何定罪，什么样的制裁才是恰当的，等等。虽然无法确定，使这一切成为可能的心理机制是如何进化而来的，但有一点可以确定，规范是需要共同遵守的。

我在几年前受邀参加了一个会议，主办方要求我以“文化是否会超越生物学”为主题做一个演讲。我在演讲时将这个题目改掉了，因为在我看来，文化不应凌驾于生物学之上。因为文化是生物学的基本组成部分之一，就像人类身体中结构独特的盆骨，以及覆盖在臼齿表面的厚厚的釉质一样。在400万年以前，我们祖先的生活中，文化所发挥的作用微乎其微。时至今日，人类已经成为一种浸润在文化中的物种，对他人所提供的信息极为依赖。

文化使人类进化出了高度精细的适应能力，彻底改变了人类的生存方式和发展方向，以及生活史。文化还使我们进化出了复杂的、具有高度合作性的、完全不同于其他生物的社会制度。这其中不存在任何非自然因素，或是非生物学因素。在过去的数百万年间，由于受到了一般意义上的进化过程的影响，人类在形态、生理和心理上发生了巨大的变化，这一切使文化得以持续进化。文化让人类成了与众不同的生物，然而毫无疑问，人类依然是大自然的一部分。

A DIFFERENT KIND OF ANIMAL

中 篇

评论篇

03
模仿、哈耶克理论和文化学习的意义

艾伦·奥尔

罗切斯特大学进化生物学教授

我不是人类学家。作为生物学家，我的主要研究方向不是人类进化，而是物种的形成和适应性。因此，我对人类文化进化的了解，从某种程度上来说，必然处于门外汉阶段。虽然我这个门外汉的观点难免会有很多疏漏，但我还是希望能给大家带来一些启发，尤其是拿人类进化与其他物种的进化做对比，找出其间的相似之处和必然的相异之处。

大体上，我赞同罗伯特·博伊德教授的观点。最重要的是，我完全赞同他所说的，人类是生物界的万物之灵。人类能够在这个星球上的任何一种生态系统中生存；至于生活方式，其他物种只有妒忌的份儿，如果它们也有妒忌之心的话。在我看来，人类的出现是进化过程中的阶段性变化之一。一些生物学家不愿承认人类的独特地位，这让我备感诧异。这可能是维多利亚时期，博物学家赫胥黎（T. H. Huxley）和其他一部分人所发起的研究项目的遗毒。那些项目非常巧妙地，有时候也没那么巧妙地从各方面将人类降格为昆虫的近亲。当然，这种观点或许是对的，但有时候所谓正确的观点可能是答非所问。

人类的生物本性是毋庸置疑的，但是为何会迥异于别的物种？不得不承认，非常重要的一点是，只有以人类的生物本性为前提，这一问题才是有意义的。在达尔文提出进化论之前，这个问题的答案很简单，而且毫无价值：因为我们更接近天使，而非动物。在进化论被提出之后，人们开始质疑以前的答案，于是，这个问题变得更加重要了。

我发现，博伊德针对这一问题的研究思路十分有趣，其吸引力主要表现在以下几个方面：首先，他认同生物学和文化学在人类进化过程中所起的作用，并对两者都非常重视，而且这种重视并非只停留在口头上。他还认同达尔文的自然选择理论所具有的解释力，不管这种选择是发生在基因层面上的，还是发生在文化层面上的。我要强调的是，对于达尔文式的文化进化论，我原则上是赞同的。

无疑，这一理念至少可以追溯到生物学家乔治·普赖斯（George Price）所提出的普赖斯方程（Price Equation，1970，1972）。普赖斯认为自然选择会同时发生在多个生物层面上：基因层面、有机体层面、物种层面，以及文化层面。进化生物学家有时候不愿意涉及较高层面的自然选择问题，并不是因为这种层面上的自然选择理论在理论逻辑上与其他层面上的有所不同，也不是因为这种理论不太好掌握，主要是因为

较高层面的自然选择行为往往缺乏确凿的实证材料，或者好不容易找到了确凿的材料，却发现用其他理论也解释得通。不过，这是一个复杂的技术性话题，无须在这里详细探讨。至于我的观点，可以参考本人在2015年发表一些的文章。

我在这里想说明的是：第一，从原理上来讲，自然选择可以作用于任何级别的生物层面；第二，如博伊德所说，即使人类文化的进化形式并不完全等同于基因或有机体的进化形式，但在这几者之间，或许存在一定程度的相似性；可以对此进行一些研究，以增进对人类文化进化过程的了解。当然，达尔文式的文化进化论还存在其他一些重要的看法，其中包括意大利种群遗传学家卢卡·卡瓦-斯福尔扎（Luca Cavalli-Sforza）和英国免疫学家马克·费尔德曼（Mark Feldman）等人所提出的基因-文化协同进化论（Gene-culture Coevolution）。

其次，博伊德的研究思路以精确的数学模型为基础。这种数学模型是种群遗传学家们再熟悉不过的了，而且具有无可估量的价值。博伊德在演讲中对自己和社会学家查尔斯·佩罗、人类学家克里斯蒂娜·莫亚（Cristina Moya）（Perrault，Moya，Boyd，2012）所构建的一个正式模型进行了分析和总结，我们就以此分析为例，来说明一下数学模型的重要性。他们探讨的问题是，在环境线索非常杂乱且不确定的情况下，个体是应该模仿他人的行为，还是应该相信基于环境线索所做出的推断。他们的结论，同时也是正式的数学模型的分析结果是：在环境发生缓慢变化，或者环境线索非常杂乱的情况下，个体应该模仿他人的行为。在我看来，这正是解决此类问题的正确途径。

当然，这种分析方法并不能给我们带来这类问题的最终答案。我认为，要想获得最终答案，可能需要做更多设计严谨且巧妙的实验。以门外汉的角度来看，或许会得出这样一个较为合理的判断，那就是在这一研究领域中，此类实验是非常滞后的。如果说种群遗传学的发展能带来一些启发的话，那么非常重要的一点就是：先有正式的理论，后有合理的实验，而且理论和实验之间的时滞较长。只有先建立起成熟的理论，才能清楚地知道需要做什么样的实验。几乎所有的种群遗传学经典实验，包括特奥多修斯·杜布赞斯基（Theodosius Dobzhansky）关于“区分平衡选择和诱变选择平衡过程”的实验，都离不开数学模型理论家们在前期所做出的贡献。这些理论家包括遗传学家塞沃尔·赖特、遗传学家罗纳德·费希尔（Ronald A. Fisher）和种群生物学家J. B. S.霍尔丹（J. B. S. Haldane）。

最后，也是最重要的一点，博伊德的研究思路有明确的边界，没有随意引申概念。虽然进化心理学或文化基因学也有不少言之有理之处，但其支持者有时候会忍不住将一个简单的概念扩大化、引申化。

除此之外，每一位科学家都应明白，有吸引力的理论并不一定就是正确的。事实上，让人感到窘迫的是，很多时候二者都背道而驰。因此，我们接下来就不再讨论宏观的方法论问题了，转而探讨博伊德理论中的两个小问题。当然，这两个小问题并不会对其理论的总体思想构成严重的威胁。

超级大脑与精准模仿

博伊德的主要观点是，人类的文化学习能力是造就人类独特地位的根本原因。总的来说，他认为：第一，从不同时空背景下的人类栖居环境类型来看，传统的基因层面上的适应理论无法解释人类为何能进化得如此之快；第二，个体学习理论对个体的学习能力要求太高。

我赞同第一点。基因变异促生适应性的速度要比许多进化生物学家曾经以为的要快一些（Lescak et al.，2015）。虽然这种适应性对晚期智人在某些方面的成功起到了重要作用，但无法解释人类为何能如此快速地扩张到地球的每一个角落。对于第二点，即个体学习能力的有限性，博伊德所引用的欧洲探险者的迷失经历很有说服力。如果人类大脑的超强能力是人类在地球上兴盛不衰的主要原因的话，那么当那些欧洲探险者发现自己迷失在一个陌生又恶劣的环境中时，也理应能够活下去才对。然而，他们却死在了那里。令人悲哀的是，个体学习能力没能帮助他们成功掌握在当地存活下去所需的本领，比如编织渔网等，而当地人在做这些事情时可谓驾轻就熟。

在演讲的最后，博伊德用大量的篇幅比较了两种解释人类独特性的理论。第一种理论，即传统理论认为，人类之所以如此聪明是因为拥有超级大脑。正如进化心理学家所指出的那样，更高的智力水平使现代智人得以进入一个开放的认知生态位（Tooby & DeVore，1987）。第二种理论，即博伊德所提出的观点认为，人类是模仿他人行为的高手。

我所关注的是，博伊德过分夸大了两种理论之间的差异。要想成功模仿他人的行为，必须具有很强大的智力，这一点非常关键。当今社会过于看重创新的重要性，因此人们很容易低估模仿所需的智力。这种智力或许是斯坦福-比奈智力量表（Stanford-Binet Intelligence Scale）或

美国学术能力测验（SAT）所无法涵盖的。实际上，精准模仿需要很高的社会性智力，因为需要解读微妙的线索，以及解决令人头疼的框架问题。博伊德承认，自己并没有考虑其中的一些要素，尤其是框架问题。比如，他提到，假如我看到你在制作弓箭，而这一过程会涉及很多行为：你盘腿而坐，削尖箭头，选择好羽毛并粘在箭尾，同时嘴里反复哼唱着小曲。很难确定对于制作弓箭来说，哪些行为是必要的，哪些是不相关的。区分相关的行为，需要极大的智慧。

至少有两种方式可以帮助我们了解在这种情况下所需要的智力情况。最著名的例子莫过于人工智能。正如这一领域的研究者长期以来所认为的那样，对"在餐馆点餐"和"编织渔网"之类的事所涉及的知识进行编码，是非常困难的。要想完成此类任务，人们就需要动用自身储备的大量隐性知识来进行判断：哪些情况在哪个层面上的知识，与眼下的情况有关或无关。当我们试图在计算机程序中对此类行为进行编码时，需要将这些隐性知识转化成显性知识。

另外，还可以通过实验的方式来观察精准模仿之于人类学习能力的难度。在这里，我们只能在头脑中将这个实验"做"一遍。现实中的实验材料可以在本段结尾处所提及的文献中找到。假设，我们试图通过大量的接触和训练，慢慢教会另一物种的某个个体完成一些复杂的任务。接下来的问题是，该物种的其他个体能否成功模仿榜样个体在完成复杂任务时所做出的行为。鉴于两方面原因，这个实验很有可能会失败。一方面，尽管我们进行了耐心的教导，但该个体的智力不足以完成这些任务；另一方面，该个体学会了如何完成任务，但其他个体的智力不足以成功模仿相应的行为，要么只能部分模仿，要么根本就无法模仿；当然也有可能缓慢地掌握这些行为，但是其速度与那些没有接触过榜样的个体无异。对多种物种所进行的多项实验表明，第二个方面的原因更多一些。关于此类实验的详情，请参阅认知科学家弗朗西斯·苏比尔（Francys Subiaul）等人在2007年所写的文章《关于猴子实验的复杂历史的回顾》。

"超级大脑"和"精准模仿"这两种理论之间的差异并没有那么大，这意味着，将进化心理学观点和博伊德的观点调和到一起是有可能的。用一个简单的完全假设性模型就可以证明这一点。设想一下，自然选择使人类得以有效模仿他人行为，而这会给潜在的创新者带来更大的选择压力，他们必须让自己所掌握的技能变得更加娴熟，也就是说，自然选择会推动人类步入认知生态位，以使创新者可以用某种搭便车者难以复制

的方式进行创新。

由此便不难理解，为什么创新者愿意接受创新的成本，同时不允许任何人无偿使用其劳动成果。但是，创新技巧的强化会反作用于自然选择，使之选择出更高明的模仿技巧。用进化心理学的话来说，我们会看到一个正循环，即复杂的创新基因组会不断带来更为复杂的模仿基因组。这便是达尔文式的进化过程，如此这般，循环往复。现在我已完整地构建出了这个模型，说实话，我并不是很看重它。据我所知，这方面的文献早就有了。我只是想特别强调一下，调和进化心理学观点和博伊德的观点的方式有很多。

哈耶克与对传统的尊敬

我所关注的这一点，可能有点出乎进化生物学家们的预料。博伊德在演讲中探讨了一个特别有趣的实证发现：在许多文化中，很多人都无法解释清楚为什么自己会恪守某些禁忌。他用了不少篇幅来列举实例，其中一个是斐济岛上的妇女需要对某些海产品忌口。简单来说，斐济岛上的妇女在怀孕或哺乳时，依照传统是不能吃海龟和其他一些海产品的，而其他人却不用对这些食物忌口。在被问及这些食物会给自己以及孩子带来何种后果时，这些妇女给出的答案千奇百怪，比如新生儿的皮肤可能会因此而变得非常粗糙。

尽管这些答案看起来荒诞不经，但这些禁忌已被证实是非常理性的，能使妇女们避开那些雪卡毒素含量较高的食物。这种毒素会在诸如海龟一类寿命较长的海洋生物和食物链顶端的海洋生物体内不断积存。正如博伊德所注意到的那样，人类摄入此类毒素后会出现呕吐、腹泻、关节疼痛，以及其他一些症状。这种毒素对胎儿和婴幼儿的危害尤为明显。通过研究，博伊德得出结论：虽然群体成员可能对遵守的规范和禁忌不甚理解，但这些规范和禁忌往往是理性的。

这不禁让人想起社会理论家、经济学家弗里德里希·哈耶克的观点。哈耶克提出过一个非常著名的论断：规范、守则和制度大多都是社会进化的产物。他所说的社会进化类似于达尔文所提出的自然选择，只不过发生在群体层面上，并有赖于持有不同道德规范的群体之间的竞争。一些规范能使群体兴盛，而另一些则会导致群体衰败。因此，我们今天所看到的规范、守则和制度并不是各种可能性中的随机产物；相反，它们代表的是造就功能健全的理性社会的事物，而这类事物绝不是随机产生

的。

哈耶克还认为，在《物种起源》一书出版前的很长一段时间里，已有一些思想家看到了这种类型的社会进化，他将这些思想家称为“达尔文之前的达尔文主义者”（Hayek，1984a），包括伯纳德·曼德维尔（Bernard Mandevill）、大卫·休谟（David Hume），以及很多苏格兰哲学家，还包括爱尔兰人埃德蒙·伯克（Edmund Burke）。伯克于1790年出版了著作《对法国大革命的反思》（*Reflections of the Revolution in France*），并在书中对突然抛弃社会进化而逐步形成的、会造成严重行为后果的规范和制度进行了深刻的反思。包括哲学家詹姆斯·西科德（James A. Secord，2008）在内的一些达尔文理论研究者认为，达尔文阅读过上述政治学家和经济学家的著作，并深受他们的影响。

哈耶克还进一步论证了，在一个充满分散的隐性知识的世界里，个体常常无法说清传统为何会以现有形式存在。事实上，在某些极端情况下，这类理解可能已超出了个体的认知范畴。一个典型事例是市场机制和价格水平。没有人能说清楚，市场上任何一件产品的价格何以能反映其价值，比如一辆崭新的吉普车。原因在于，吉普车的价格会受到诸多错综复杂的因素的影响，而这些因素的互动方式可谓神秘莫测。无论是钢铁的供给，还是橡胶生产国的工人动乱，无论是打车软件的出现，还是计算机行业对塑料的需求，以及标准普尔500指数的表现，莫不如是。

然而，市场会以一种看似神奇的方式对这些因素做出权衡，并以出清价格来售卖产品。由此可见，市场机制作为一种社会制度，为人类提供了一种解决方案，而这种方案是任何理性的个体或计划制订者都无法提供的。

在哈耶克看来，人们必须“尊敬”各种社会规范和社会传统。不管能否对其做出合理的解释，这些传统和规范都是经过时间检验的社会进化的产物。至少，它们使人类社会得以延续至今，还让人类过得不错；至少，它们在历史的长河中战胜了其他一些种群所支持的很多传统和规范。

哈耶克还有个不太为人熟知的观点。他认为社会科学要想具有合法性，就必须深入了解传统中所蕴含的非人文知识。他对科学哲学进行过很多论述，譬如于1952年出版的著作《科学的反革命》（*The Counter-Revolution of Science*），以及收录在《哈耶克文集》（*The Essence of Hayek*，1984b）中的一系列文章。他在论述中指出，虽然自然科学验

证了理性的力量，但是从社会科学的角度出发，必须谦虚地承认，理性是有局限性的：传统行为背后的复杂因素可能是任何人类个体都无法理解的。因此，只有克服他所说的“知识的伪装”，或者说“傲慢的假设”，理性的个体才能设计出比文化进化所带来的非人文解决方案更优秀的方案。

现在，博伊德应该很清楚我所说的话了。事实上，他和彼得·里克森于1993年所发表的文章《理性、模仿和传统》（*Rationality, Imitation, and Tradition*）对哈耶克的思想已有涉及。当然，不是说博伊德的观点和哈耶克的观点是一致的。显然，它们并不一致。博伊德的观点更具描述性：人类是模仿他人的高手。哈耶克的观点则具有一定的规范性：人类应该做模仿他人的高手。这也就是说，不管能否理解，人们都应该尊重传统。而且，在文化进化的具体形式方面，博伊德的理解和哈耶克也不一样。与博伊德不同，哈耶克对“有利的突变”，即成功的创新是如何产生的，总是含糊其词。是因为某些个体天生就更聪慧吗？还是因为社会在以某种形式进行着随机实验？另外，哈耶克很多时候喜欢用散文的笔调来表达观点，因此难免会有些含糊不清，虽然文笔很好；而博伊德偏好用密密麻麻的数学模型来表达观点，譬如他和里克森于1993年所发表的那篇文章。

这两种观点尽管来源不同，但结论的确很相似。我在想，我们这些科学家对哈耶克的了解是否有所欠缺，生物学家尤其应该感到羞愧。关于这个问题，请参阅我在《达尔文与达尔文主义》（*Darwin and Darwinism*）中所做的详细阐述。我们之所以对哈耶克不够了解，是因为与所研究的主题有关。学者们之所以会忽视哈耶克，部分原因在于，从政治角度来看，他的观点在学术圈里已经没那么流行了。毕竟，人们已经对哈耶克形成了这样一种印象：他的观点主要是在为传统的政治秩序做辩护。很多人都对此感到不适。不过不要忘了，哈耶克还写过《为什么说我不是一个保守主义者》（*Why I Am Not a Conservation*）一文。我们应该改变对他的刻板印象。

最后，我们或许会发现自己不但做了些徒劳无功的事，而且还被困在一个奇怪的位置上。因为，如果真像哈耶克所认为的那样，进化思想萌芽于那些研究人类史前文化和历史文化变迁的社会科学，那么我们现在所做的一切，岂不是把进化思想又反哺到了其肇始领域？所建构的清晰的人类文化进化模型，岂不是把哈耶克口中的那些“达尔文之前的达尔文主义者”再次达尔文化了？

04
适应不需要洞察力？

金·斯特林

澳大利亚哲学家

对这次演讲所涉及的主题，我一直以来都很认可。我特别赞同罗伯特·博伊德所阐释的其研究项目中的四个核心观点：

1. 的确如博伊德所描述的那样，人类，主要是近代人种是一种独特的生物。以生物进化的标准来看，人类这一物种脱颖而出的时间虽不长，但速度却很快。因此，我们不仅需要注意人类和相近物种之间的巨大差异，还要注意人类脱颖而出的惊人速度。

2. 人类不仅数量众多，而且遍布世界各地。这是因为人类能够在错综复杂的环境中开采资源，并应对各种危机。这种环境适应能力主要体现在信息和行为等方面，而非生理方面。当然，文化学习在其中也起到了一定的调节作用。因此，在解释人类与相近物种以及祖先之间的差异时，社会学习，尤其是累积性的社会学习的进化是一个非常关键的因素。这种区域性的适应能力的形成需要经历许多世代，是一个种群层面上的过程。人类个体的适应能力、表征和理解周围环境的能力，以及表征和应对未来突发事件的能力都是相当独特的。然而，个体层面的认知能力并不足以解释人类的人口规模、生态系统和地理分布，以及在社会层面与技术层面上对特定环境的适应能力。我认为这些能力是必不可少的，虽然博伊德对此既赞同又不赞同。不过，仅仅拥有这些能力是远远不够的。

3. 只依赖地方性信息是不够的。人类社会生活的规模、范围，以及生态多样性还取决于社会合作。社会合作在其他哺乳类动物中非常少见，但对于人类社会而言，社会合作已达到了史无前例的水平。人类的合作具有强制性，因而博伊德认为，需要对人类的合作进行独立阐述，在这一点上他是对的。即使是小规模的人类社会也不是完全由近亲构成，而且从来都不可能是。因此，用那些解释亲缘性合作行为，比如群居类昆虫的合作行为的进化模型来解释人类的合作行为是行不通的；作为概括性论述的直接互惠理论也是行不通的。一旦合作人数达到三至四人及以上，因失败而放弃合作的行为便会对维持合作稳定性造成极为不利的影响。因而，人类的合作行为的确需要详加解

释。

4. 除了这些实体论之外，博伊德和彼得·克里森还找到了恰当的方法论来考察人类群体生活的形成过程。通过对民族志和考古学资料进行对比和分析，围绕人类个体的动机与认知进行实验研究，以及对这些数据流和模型的联系进行分析，我们对这方面的知识有了更深入的理解。这些模型有助于我们理解，人类的决策与行为模式是如何在群体层面上不断升级的，而群体又是如何随着时间的推移而进化的。

对于博伊德的观点，我没有太多的批判，主要关注的是我俩在某些观点上的细微差别，以及这些差别会带来的问题。我的评论将围绕以下三个议题展开：个体智力在社会学习中所扮演的角色、对模型的理解、将博伊德的理论扩展到不同的时期和社会学习领域。

个体与社会学习

我同意这个观点，人类所具有的独特智慧并不足以解释人类的生态灵活性和适应能力。基于不少证据，博伊德怀疑，人类的智慧并非必要因素，至少在解释各种文化的变量传播时如此。这与解释文化起源大为不同。[1]在解释累积性的社会学习对塑造环境适应性反应所起的作用时，博伊德这样写道：

> 但在我看来，人们通常会这样想：创新是很难的。发现大柄苹有毒并不是一件容易的事，而找到去除毒素的方法更是难上加难。好在，确定一种行为是否有益相对比较容易。因此，一种发明一旦出现，只要被确认是有益的，便会被传播开来。

博伊德将这一现象称作“图书馆”模型，并引用人类学家约瑟夫·亨里奇（Henrich，2010）关于斐济人食物禁忌的例子对这一共识进行了论证，接着又以房屋建造方面的例子为佐证。斐济人禁止怀孕的妇女食用处于食物链顶端的海洋生物，以防止她们中毒。这一禁忌被赋予了道德意义，而不只是一个慎重的建议。亨里奇的民族志表明，尽管恪守这种食物禁忌之人并不清楚为什么要这样做，但在其生活实践中，这种做法是长期稳定存在的。虽然长辈们解释不清这些禁忌食物的危险性，但这种生活实践方面的建议大多来自他们，而非“自食其果”之人的现身说

法。

由此可见，适应性行为可以通过文化进化的方式稳定下来，甚至被创造出来，即使人们无从知晓这种行为的益处和影响。人类学家马尔温·哈里斯（Marvin Harris）曾发起过一个研究项目，从隐性适应的角度解释禁忌和食物偏好，尤其是那些出人意料的食物偏好（Harris，1985）。然而，亨里奇和博伊德却构建出了一个清晰的模型来解释此类隐性适应行为。

博伊德认为，行为主体在选择榜样和信任榜样方面是有策略的，因此，社会学习才成了一种建立在可靠信息基础之上的学习。然而，人类的智慧大部分都体现在对榜样的选择上，而不是对榜样所发出的信息流进行理解。尽管这种理解具有降噪的作用，但并非至关重要。我将再花一些时间来阐述这种图书馆模型。第一，我要强调的一个事实是，探究某种做法为何会具有适应性是一个渐进的过程；第二，理解这个过程通常需要一些“默契”，因为它不是明确的、陈述性的、说明性的知识。

博伊德所列举的斐济民居的例子并不足以解释，为何未采用标准化设计的房屋会在飓风中毁于一旦，而且也并不能证明房屋建造者不知道如何进一步加固房屋。虽然缺乏显性知识，但这并不能作为他们不具备隐性诀窍的证据。这个问题很关键，但博伊德等人却没有对此进行相关调查研究。因此，在充分了解了前因后果之后所做出的实践，与那种基于信任的盲目模仿是极为不同的。实际情况很有可能是，由于社会交往对社会实践中智力水平的依赖程度不同，因此某种知识的社会传播范围也会不同。然而，我个人的直觉是，技能的传播在大多数情形下都起到了关键作用。

在食物禁忌这个例子中，博伊德重构了和社会传播有关的“智力-理解”模型，就此例而言，该模型所预测的社会动态是这样的：

> 那么，为什么斐济人会相信父母、亲戚，以及有智慧的女性所告之的这些经验呢？依照“文化是图书馆”这个观点，答案如下：这些学习对象会给出极好的理由让他们相信孕妇和哺乳期妇女不应该吃海鳝、海龟等食物，并遵照这些信仰行事。这些信仰需要具备足够的说服力，只有这样才能让一个品尝过美味海龟的妇女，在其怀孕时顾及这些食物给未出生的孩子所带来的不利影响，从而选择天天吃煮熟的木薯。这些信仰必须具有强大的说服力，只有这样人们才敢指责一个吃海龟的孕妇，

> 即便这种指责可能会冒犯她和她的亲属。食物和疾病之间的因果关系将会得到强化，因为事实证明这种关系是真实存在的。例如，村子里的某个人注意到，某位妇女在不知道自己已怀孕的情况下吃了不少海龟，后来生出的孩子在发育方面确实存在一些问题。

博伊德的人种志表明，此类事件并没有发生过。这种行为规范的传播有赖于一个更简单、更容易让人相信的启发式社会学习策略。

这个策略有多典型？可以参考一下通过示范和实践来完成的技能传播。不同于仅在观察条件下所展现的技能，在传播中，技能的示范要求行为主体展示出行为的因果关系。我虽然不是一个碎石匠，但体验过几次这方面的课程。我师傅对这项技能的步骤是非常了解的，他需要将这些步骤依次展示出来，做出说明，并就击石的角度、原料的选择、岩石的几何形状，以及实现目标的顺序给出建议。考古学家彼得·希斯科克（Peter Hiscock）认为，传授碎石技能所采用的主动教学法可能非常古老，因为无人指导的试错学习法是极其危险的，对此我深信不疑。岩石碎片非常锋利，错误的敲击会让碎片四处飞溅，并导致无法预料的严重后果（Hiscock，2014）。

与此类似，生态学家路易斯·利本伯格（Louis Liebenberg）曾对非洲南部狩猎者高超的猎物追踪本领进行过研究，并对所需的认知能力进行了分析。这些狩猎者在对猎物进行追踪时运用了大量的图形再认知技能，并且对大量自然历史信息进行了细致入微的观察（Liebenberg，1990，2008，2013）。这些狩猎者不仅知道哪些踪迹属于哪只动物，还知道同一只动物在疲劳、惊恐、紧张或放松时的不同行踪。这些技能的社会传播有赖于学习者的智力水平：

> 青少年和年轻人会观察部落中最优秀的狩猎者是如何谈论与解释动物足迹的。根据这些足迹，高明的追踪者可以推断出动物的年龄、性别、身体状况、疲劳程度等信息。要掌握这样的技能，需要对特定物种的习性、饮食习惯、社会构成，以及日常行为模式等均有一定的了解（Henrich，2016）。

狩猎社会的经济结构要求人们掌握和领地有关的丰富的自然历史知识和地理知识。博伊德在讲述探险者罗伯特·伯克和威廉·威尔斯所经历的悲壮又颇具启示意义的探险故事时这样写道：“一部关于扬德鲁万塔人

的‘自然历史书’可能会有上百页，并分为‘如何游戏’、‘有效的打猎技术’、‘如何找到水源’和‘如何处理有毒的蕨类、薯类和苏铁类植物’等板块。”他们的后代无须通过个体学习来获得这些知识，社会学习就已足够了。

不过，后代们会通过实践来不断验证和完善这些信息。他们会不断地用外界信息来对技巧、技术以及地方性知识进行检验。如果所获得的外界信息已经过时，或者在传播过程中失实，那么就需要用一些相对直接的方法来纠正错误。一个地区的自然历史和微观地理并不会一成不变，小溪会改变流向，旧的水坑会发生淤塞，新的水坑会出现，动物的习性和数量会发生变化。狩猎者会从同辈和长辈那里了解到这些知识，同时也会通过观察外界来进行了解（Hewlett et al.，2011）。人们会接受规范与禁忌之类的法则，并依照这些法则“盲目”行事。如果某种禁忌能够帮助某个妇女避免某种危险，而这种危险在过去从未真正出现过，那么对于她来说，错误信号是极其微弱和模糊的。有人说这些技能和信息在物质世界与生物界中均发挥着指导性的干预作用，这并非事实。这是我在2012年所关注的主要议题之一。

在我看来，这种食物禁忌在两个方面上与一系列重要的事实产生了冲突。第一，这个例子是纯粹的社会传播实例，除了少数人有过亲身体验之外，恪守这种禁忌的大多数人所接受的信息，也就是处于食物链顶端的海洋生物很危险这个信息，来自社会学习。技能的传播涉及复杂的社会学习过程，且社会信息的输入至关重要。但是，仅依靠口头指导或行为示范，人们是不太可能学会如何追踪猎物或制作手斧的。只有在社会性的指导与周围世界发生联系时，个体才有可能获得相应的技能；行为主体往往会在实践中进行学习，不过其学习方式是由社会环境造就的。虽然信息来自物质世界，但在榜样、示范、建议、支持，以及成功和失败的作用下，运用一些技术性的专业词汇便能表达出原材料之间的微妙差异（Stout，2002）。就这一领域而言，社会层面上的信息与个体层面上的信息必然会不断结合。

第二，斐济人的食物禁忌是文化适应方面的典型实例，但算不上是累积性文化适应方面的实例。累积性文化适应是适应能力的一种，比如对火的使用。这类技能超出了个体学习的范畴，需要通过渐进提升来获得，而且很可能跨越了数代人。地方性文化适应的作用及其多样性，在很大程度上可以用不同环境中的累积性反应来解释，而食物禁忌这个例子并不能证明这一点。博伊德所列举的动物筑巢的例子也并不能证明，

规则简单的社会学习能够促成复杂的社会适应能力。这些行为也不是以学习为中介的适应性行为，而是一种不了解因果关系的适应性行为；是无须理解因果关系的学习行为，而不是高度真实的社会学习行为。

正如我在前文中所说，能否知道某种做法为何有效，与理解程度密切相关。为何某些形状的岩石可以被制作成有用的工具，而其他的则不能？许多技艺高超的碎石专家都对所用材料的科学知识知之甚少，但对原材料的特性了如指掌，比如从某个要求严苛的角度进行敲击会出现某种状况，但这种知识所适用的范围和情况是有限的。

我估计，生态学家路易斯·利本伯格谈到的追踪技能或许也是如此。那些非洲狩猎者深知动物在基底层上留下痕迹的过程，而这种知识和动物的数量、基底层的情况以及环境的情况都有关。他们的专业知识可能无法应用于冰天雪地，而高纬度地区的狩猎者定会拥有“雪地”技能。即使是斐济人也不会完全盲目行事。正如博伊德在报告中所说，斐济人相信，违反食物禁忌很可能会致病，并普遍认为吃鱼会生病，而且有些鱼比其他鱼更危险。这种普遍认知增强了禁忌的可信度，而且与其所拥有的和世界运行机制有关的一般经验高度吻合。

图书馆模型并不正确。社会学习策略评判的是行为主体，而不是行为主体所展示出来的文化变量，因此才能对地方性的适应反应起到支撑作用。然而，如果在一个模型中，行为主体被刻画成“会产生反射性冲动，并会盲目模仿他人”的形象，那么这个模型一定是错误的，尽管被模仿者可能是经过精心挑选的榜样，但这种冲动也只有在外界信息足够明确时才会起作用。特别是这一模型低估了混合学习的作用，没有注意到学习主体在学习过程中并不需要在社会信息和直接信息之间二选一，而需要将二者结合起来。在我看来，在传播技能的过程中，混合学习尤为重要。因此，在阐述与文化变量有关的社会传播需要哪些认知能力时，应该讲清楚，认知需求与文化变量的本质及复杂性之间的关联方式出现了哪些变化。

对博伊德构建的模型的分析

博伊德用了一个和社会学习有关的简单的进化动力学模型来支撑其人种学理论。在被他称为“盲目模仿的进化”的章节中，他对这一模型进行了描述，并认为该模型可以证明自然选择更青睐这样的心理：人们会接受某种观念仅因为他人也持有这种观念，以及行为主体具有“采纳他人

想法的倾向”。但是，是否有必要这样来解读这一模型的结果，我并不确定。该模型探讨了复杂环境中的社会学习驱动机制。行为后果取决于行为主体在特定环境状态下所做出的行为选择。行为主体掌握了来自外界的直接信息，然而这些信息很杂乱，因此他们并不能明确其他人会作何选择，毕竟周围人所掌握的外界信息也同样杂乱。这一点也不奇怪，直接信息越杂乱，想要直接了解外界的难度就越大，行为主体也就会越依赖间接的社会线索。

对我来说，这是一个既简单又智能的策略性学习模型。行为主体会进行自我评估，如果认为自己所掌握的外界信息是可靠的，就会使用这些信息；如果认为信息不够可靠，就会利用信息流的优势，去观察群体中其他人的反应。[2]这是他们的最佳选择。如果行为主体在个体学习和社会学习之间做出了最佳的行为选择，那么在解释这种反应时，就不需要以“采纳他人想法”的心理倾向为前提了。事实上，在这个经过简化的模型中，行为主体尽可能地做出了恰当的选择。想要验证“采纳他人想法”的心理倾向的确存在，就要找出和个体接受当地主流观念有关的人种学证据，即使有更可靠的证据验证了其他的评估结果。此外，还必须考虑信息成本，以及作为特定环境中的特立独行者可能要付出的声誉上的代价。

显然，宗教可以给我们提供许多例子来证明：行为主体会忽视外界信号，重视社会信息。但在宗教领域中，特立独行者同样要付出高昂代价，也就是说，我们可能会惊讶地发现，偏好社会信息是存在隐性风险的。在这方面，我认为，人类学家奥利维尔·莫林（Olivier Morin）的看法很正确。他指出，许多社会信息都只是昙花一现：谁在哪里看到了什么。那些广泛分布于社会网络中的信息，在公共信息领域里不过是些小样本而已，而且还是非随机性的。与他人言行有关的信息也会对我们产生影响，因此我们需要过滤掉大量的无关信息（Morin，2016）。

博伊德用材料和模型支撑了如下观点：社会学习在很大程度上属于简单的遵照惯例的启发式策略，且能在许多环境中发挥作用。对于行为主体而言，鲜有人能够以较低成本找到可靠又直接的环境信息样本。因此，人类进化出了一种学习机制，用来权衡社会学习的成本与收益。人们通常偏好相对简单的社会学习规则，但在发现自己可以轻松便利地从所处环境中获取信息时，就会开始关注环境信息的有效性，进而克服这种偏好。

莫林在其著作《传统是如何生存和消亡的》（*How Traditions Live*

and Die）中，言辞激烈地驳斥了上述有关社会学习的观点。他认为社会学习虽然重要，但具有策略性，而且关键取决于行为主体对自身所处的信息环境，以及潜在的模仿对象所处的信息环境的评估结果；人们实则不会也不应该使用这种被简单化且墨守成规的启发式策略（Morin，2016）。

到目前为止，莫林对博伊德的批评主要针对的是，后者对模型的描述过于简单。这样的批评显然是无的放矢。现实环境中的主体行为远复杂于博伊德所构建的模型中的主体行为，而构建模型的初衷之一即是删减掉不必要的信息。

关键在于，在种群动力学上，这两种主体行为之间的差异是否具有某种影响力？至少有三种观点可供探讨，而且只能借助形式模型来探讨，因为形式模型能够表征简单策略化的社会学习者与莫林所说的高度策略化的社会学习者之间的差异。第一种观点是，这种差异并不具有影响力。因为在由信任型社会学习者构成的人群和由辨识型社会学习者构成的人群之间，存在相似的驱动机制；对比人种志和考古学研究所提供的解决方案即可见，这两种机制是相似的。第二种观点是，在具有策略性和辨识能力的文化学习群体中，累积性的文化学习是非常少见的，甚至不存在。如果个体对自身的判断力过于自信，那么就会令当前可信任的认知资本（cognitive captital）[6]水平难以得到维持，从而导致下一代无法在当前基础之上更进一步。对于累积性的社会学习而言，高度的社会信任是必要条件之一，人种志和考古学所提供的渐进性适应能力方面的证据表明，在传统社会中，人们更信赖社会学习。第三种观点是，模型表明，作为累积性学习的工具，辨识型社会学习的功能性要比信任型社会学习的更为强大。在以辨识型社会学习为主的群体中会出现更丰富的微小创新，而且这些创新是可信任、辨识和接纳的，同时还能更高效地摒除错误。模型还表明，辨识型社会学习者通常不会过分追求局部最优化，不会陷入无用的纳什均衡（Nash Equilibrum）[7]陷阱之中，而且面对环境变迁也很少会感到不知所措。

由此可见，在经验证据方面，博伊德的论述和莫林的论述的确存在差异。如果人类都是信任型社会学习者，那么考古学和人种志的相关资料是否能证明我们所预测的“适应失败”现象是广泛存在的，或是只存在于某个特定领域、某个具体环境中？是否能证明存在大量的偏离局部最优化的现象？如果没有形式模型，就没有办法弄明白两人的观点是否真的存在差异，且这种差异是否会体现在经验证据方面。可以看到，博伊

德和彼得·里克森为方法论所做出的贡献是十分重要的：综合使用了语言分析、人种志资料和实验材料以及形式模型。当然，这并不意味着他们会放弃选择。第四种观点是由认知学家罗恩·普拉纳（Ron Planer）和丹尼尔·丹尼特（Daniel Dennett）[8]为我提供的：或许，社会学习的启发式策略原本很简单，但后来在进化过程中逐渐变得微妙了。

关于规范的不同思考

博伊德在演讲中较少提及的一个论题是：人们能在多大程度上，对社会学习策略在不同领域和不同时间内所发生的变异，做出何种程度上的预测。

考古学资料显示出了一个非常突出的特征：技能经由漫长且极度缓慢的历史进程而得以累积。从大约350万年以前到大约40万年以前的这段时期内，除了旧石器时代的技术之外，能从资料中看到的新技能只有手斧的制造和火的使用。此后30万年间，出现了新石器技术和复合型工具。从大约10万年前起，各种技能开始加速出现。直到最近，制造、储存和累积的认知资本才出现了明显的变化，这种变化可以用个体基因差异所致的个体认知能力差异来解释。人类学家萨莉·麦克布雷迪（Sally McBrearty）和艾利森·布鲁克斯（Allison Brooks）于2000年发表了具有里程碑意义的论文《那不是一场革命：对现在人类行为起源的新诠释》（*The revolution that wasn't: A new interpretation of the origin of modern human behavior*），此后，对此类变化的解释就被强调社会组织差异或社会网络规模差异的理论替代了（Henrich，2004；Powell et al.，2009）。

彼得·里克森和博伊德就是这一转型时期的重要代表（Richerson Boyd，2013）。虽然他们尚未进行过系统的论证，但我们可以从相关资料中看到一个与其研究项目相当契合的解释：社会学习的技术及工具的文化进化。通过社会化的方式来学习如何改进社会化学习的方式，社会心理学家西莉亚·海斯（Celia Heyes）着重强调了这一观点（Heyes，2012）。

因此，不同时期的社会学习特征及可靠性并非一成不变；不同领域中的社会学习也同样会发生变化。毋庸质疑的是，大体上，累积性的文化进化并不能有效地发展出民间医学，当然也有极少数例外情况发生。大部分的民间医学都很令人生畏，与此有关的极其骇人听闻的例子有很

多，有关这一点可以参考人类学家罗伯特·埃杰顿（Robert B. Edgerton）的著作《病态社会》（*Sick Societies*）。

致病环境是不稳定的。这种不稳定性的存在并不是因为病原体会快速进化，而是因为病原体的特性和毒性会随着人口规模、经济活动和人口迁徙方式等方面的变化而变化。最为明显的是，当生活状态从不断运行转变为久坐不动之后，人体就会遭遇新的病原体，如同生活在垃圾堆和排泄物周围一样。但是，人口规模和人际交往方式的确都发生了变化。这些因素决定了病原体传播到新宿主身上的速度，以及是否会在前宿主的免疫机制启动，或者前宿主死亡之前转移到新宿主身上。前宿主的死亡会导致病原体在局部范围内被灭绝（Ewald，1994）。农业生产活动的变化可能会带来宜于疟疾滋生的外部环境条件，这正是人类学家约翰·奥德林-斯米（John Odling-Smee）等人于2003年谈到的有关生态位构建的典型案例之一。同时，医疗环境却没有出现快速变化，个中情况比较复杂，且存在一段很长的过渡期。即使一些医疗措施并没有起作用，甚至还具有危害性，但许多患者依然得以康复；当然，另一部分患者未能幸免。但不管怎么说，医疗干预还是提高了人类的生存概率。

在现代社会早期，与医疗技术相比，武器制造技术拥有更丰富的专业积累。据我推测，这是因为人类选择杀戮的动机远强于治病救人的动机。当然，在拥有了武器制造技术之后，成功的创新，以及对成功的认识就变得简单多了。一项技术能否成为成功的创新，往往会很快显现出来。例如，如何用较少力气打磨边角，如何使边缘更长一些，如何将黏合剂固定在准确的位置上，反之亦然。

此外，日常技术在应用过程中会被多次测试。如果在部分群体中出现了变异（假设只要存在创新，就会出现变异），那么我们就能得到一些可做比较的数据。[3]通常，与技能创新有关的因果链相对比较短，也比较简单。这种因果链会将工具的使用与使用结果关联起来，从而让行为主体得到大量与各种使用结果有关的信息。不过这些信息非常简略，而且是概括性的。博伊德所提及的斐济人建造房屋的例子即证明了这一点。房屋建造者能很快发现房屋是否结实，能否抵挡风雨，不过可能需要花费数十年的时间才能知道房屋能否抵挡暴风雨。

我认为，在医疗技术和狩猎技能之间，对规范及规范性创新的评价可能发挥着一定的中介作用。我们假设当行为主体并不太重视社会稳定与双赢合作时，此类规范的建立就具有了可能性；但实际情况往往是，我们很难对某两种规范所带来的社会影响进行比较。不妨参考一下人类

学家迈克尔·阿尔瓦尔德（Michael Alvard）曾提及的一个经典实例：在成功捕获鲸之后，人们是如何建立起一套精巧的分配规则的（Alvard，Nolin，2002）。当事人或许并不太了解与社会稳定、社会合作、捕鲸装备和技术的投入等方面有关的其他规则会产生何种效用。分配规则的改变所带来的影响是慢慢呈现出来的，有时候甚至需要经历一代或数代人，而且这种改变时而会与其他改变相互交错。另外，当规则被打破或发生变化时，群体成员之间或许会发生争执与不和谐。因此，他们可能很清楚，在收益分配规则遭到破坏时要付出怎样的代价，但不会对规则调整的结果进行评估。

或许正因如此，博伊德及其合作者才会倾向于将规范性进化视为任意群体内部皆会发生的一个过程。就已知的适应性规范集合而言，确实是群体选择的结果。我同意以下看法：人类的文化群体很可能会成为达尔文式的种群，而且其进化过程会受到文化的群体选择的影响。然而，这种复合种群的人口规模会非常小，而且种群层面上的代际转变速度会相当缓慢，同时，如果规范以集合或系统的形式存在，那么规范的变化就会受到一定的限制。所以，我们不应对这一过程的有效性抱有太高期望。

依照这些假设，文化的群体选择是一种高明的手段，因此规范集合也就不可能十分奏效。如果规范集合比那种杂乱且无效的选择机器（尽管我一点儿也不知道该如何证明这一点）更管用的话，那么我们就可以用规范产生过程中的亲社会行为的偏差[9]来对如何维持合作性社会活动进行解释。博伊德和彼得·里克森的模型，以及人类历史都已证明，一些令人震惊的破坏性规范能够经由惩罚和内化而稳定长存。

对此，有研究者提出了一些貌似合理的解释，认为如果规范符合人们的同理心、同情心和公正感等社会情绪的话，那么其传播性和稳定性就会变得更加容易实现（Nichols，2004）。若真如此，那么在许多规范集合中，有助于促进合作的特征就应是偏差相互强化的结果，尽管其相互作用力有可能非常微弱。在这种情况下，地域性群体内部便更容易达成合作，群体也就更容易存活下去，并开枝散叶。

接下来，我想谈谈与规范有关的最后一点想法。我们可以从两个方面来思考规范及规范心理。一方面，阿尔瓦尔德所提及的分配鲸的规范是相当复杂的，需要明确地说出来，让参与者都有所了解，同时也要明确地告知参与者，这些规范都属于地方性知识。对于违反规范的惩罚措施也应该是明确、公开、可告知的，且是非常准确的；另外，这些规范

本身亦属于该群体规范集合的一部分。另一方面，此类规范并不太具有认知复杂性。譬如所谓的抱怨规范，前提是他人未履行我们所期望的职责，比如，当火快熄灭时，我们希望同伴关掉火，但他们并没有那么做，而我们的反应可能是生气、反对和敌视等，简而言之就是抱怨。博伊德所提到的有关规范的例子，以及对“规范如何维持合作”所做出的分析（即使在一个没有制度的社会中），都是比较宽泛的。

图尔卡纳人的规范是显性的。然而，从某种意义而言，显性规范并不足以构成人类独特合作行为的全部基础。例如，语言是人类合作的产物，而非先决条件。还需要进一步证明的是，博伊德和里克森对“规范在人类合作中所起作用”所做的分析，是否适用于抱怨规范或社交心理学。在人类出现独特的合作行为之前，该理论或许是合理的。

关于这一点，不妨来看看博伊德所提出的高阶搭便车行为的解决方案：对那些未对违反初阶规则者进行惩罚的人实施制裁。他提出，对违规者施以恰当的惩罚本就是规范体系的一部分。促使人们遵守初阶规范的心理机制与社会机制，同样适用于惩罚违规者，以及当有关惩罚的规范对恰当的惩罚义务不起作用时。而二阶搭便车行为及其惩罚并不具有此种独特现象，需要其自身做出解释。[4]对于图尔卡纳人所拥有的这类清晰的规范而言，这是事实；但对于抱怨规范而言，这绝非事实。并不存在什么自动化机制可以确保那些会因自身期望落空而满腹抱怨的行为主体，在他人期望落空并表现出敌意时，也会表现出不满。与图尔卡纳人的规范相反，这一阶段中并没有明确且详细的制度，行为主体会将抱怨规范当作一个整体来接受。默认模型（default pattern）以及基于默认模型之上的期望，只会零星地出现。

以进化的角度看待人类社会生活

上述讨论并非是对博伊德演讲中的核心思想的质疑，也不是对其所列举的那些示范与样本的质疑。正如我在篇头所写，我赞同博伊德研究项目中的大部分核心内容。从某种程度上来讲，这种观念上的一致性是聚合思维的结果，并在很大程度上证明了，博伊德和里克森正在尝试引导人们培养进化观念，[5]引导人们从进化的角度去看待人类社会生活的独特性，以及支撑起这种生活的能力。在提出上述问题的过程中，我自己也在尝试着证明他们的研究是具有启示性、创造性，且易于理解的。之所以提出上述问题，不仅因为它们很重要，还因为它们有助于拓宽我

们的思路。[6]

05
文化进化中的推论与假设检验

鲁思·梅斯

英国伦敦大学生物学教授

能成为评论人，我感到非常高兴。罗伯特·博伊德和彼得·里克森于1985年出版的著作《文化与进化过程》（*Culture and the Evolutionary Process*）对我学术生涯的开启影响深远。真正让我下定决心，并明确研究方向的著作是生物学家理查德·道金斯（Richard Dawkins）[10]的《自私的基因》（*The Selfish Gene*）。最初接触到这本书时，我还是个十几岁的少年，它激发了我对社会进化的思考，让我最终决定去牛津大学攻读动物学。自那时起，直到现在，动物学都是适应主义者的大本营，提倡各种适应性理论。多年以后，我读到了《文化与进化过程》这本书。我不是很擅长数学，因此跳过了书中的方程式，而那些内容几乎占据了一半的篇幅，但是主要思想对我还是很有启发的。

后来，我到了伦敦，成了一名人类学研究者。当时，和人类进化行为有关的研究领域大体上可分为三个学派：进化心理学、人类行为生态学和文化进化理论。一般来说，知道了一个人的学科背景、研究兴趣和研究方法，就能大致推断出其所属学派。当然，后来这些学派的划分变得不那么清晰了。就我而言，我认为自己的研究领域是人类行为生态学。

行为生态学是一门兼容并蓄的学科，其重要传统之一是强调对进化假设的实证检验。不管哪个科学或领域的方法与策略，只要有助于解决研究难题，行为生态学就会借用。与进化心理学家不同，人类行为生态学家从来不认为这个世界存在普适性的人性，因为人类行为生态学研究的是人类行为的多样性，对行为的基因根源或文化根源并不特别在意，而更关心与适应性有关的问题。文化进化理论关注的是文化学习及其传播机制的模型构建，以及模型之于文化进化的意义，而人类行为生态学关注的是现实世界中的行为模式及其间的关联性。

然而，正如博伊德所说，生态和文化皆是行为的决定性因素，要将二者区分开是很不现实的，因为文化是人们了解所处生态体系的一种方式，也是了解各种环境中最佳作用因素的一种方式。随着对行为起源的研究越来越深入，这两个学派开始合流，研究者开始利用各种方法，去真正考量文化传播和生态多样性的影响力，进而探究进化过程是如何激发出人类某些特定行为的。因此，时至今日，我依然很难对这两个学派做出明确的区分（Mace，2014）。

规范的制约

尽管博伊德在讨论行为进化时所秉承的学科传统与我的有所不同，但对其在演讲中所表达的所有观点，我几乎都赞同。毫无疑问，规范和准则在引导人类行为方面有着超乎寻常的重要性。作为人类学研究者，当我第一次对一个小规模社会进行田野调查时，当地的规范和准则给我带来了极大的困扰。我研究的是肯尼亚北部的嘎巴拉（Gabbra）骆驼游牧民族的生育制度。

作为一名外来者，我理应像所有未婚人士一样躺在帐篷外边，背靠一个垫子或动物毛皮，然后仰望星空。考虑到当地炎热的气候，以及肯尼亚北部迷人的夜空，户外定然是绝佳的睡眠场所。但我很快意识到，侧着身子，两腿并拢，蜷缩起来，上身套一件长恤衫，是非常重要的。入睡肯定不成问题，但我一直很担心，早上醒来时我自己向后或向前滚，然后滚到一个完全不适合女性睡觉的地方。这只是众多令人焦虑的当地规范中的一个，若是一不留神违了规，就会冒犯当地人。因为对所属社会的文化太过熟悉，所以人们并不总能理解来自其他社会的人在适应这些文化时所遇到的困扰。

我曾听说这样一件事。我所在的伦敦研究团队的成员们在一场学术研讨会之后出去喝酒，聊天时谈到了乐坛传奇人物、雷鬼音乐大师鲍勃·马利（Bob Marley）。如果你是英国人、美国人、牙买加人，甚至是埃塞俄比亚人，那么应该知道鲍勃·马利是谁。但是，如果你是中国人，那么可能从来没有听说过他。当一位新来的中国同事问起“鲍勃·马利是谁”时，其他人对她报以了善意的质疑。她感到极为尴尬，因此在又一次研讨会结束后，没有和大家一起出去喝酒。几个月后她才告诉我这件事，我感到很震惊，因为我知道，即便其他人主观上并没有想要给她施加压力，甚至根本没有意识到这方面的问题，但这位女士已然感受到了极大的社会压力。这即是已被社会化的规范与准则。想一想那些跨文化背景的，围绕婚姻、贸易、政治处境，以及冲突等各方面问题所进行的谈判有多么复杂，就会明白人们确实受到了规范与准则的制约。

我们已经知道规范的制约作用很明显，不过还有一些问题尚待回答，比如这些规范是如何形成的，不同地方的规范有何差异，规范是如何得到强化的，以及变革规范是难是易，等等。之所以要关注文化传播机制，其中一个非常重要的原因就是，我们想知道这些传播机制能否从其他某些行为发展和进化机制中衍生出完全不同的结果。文化群体选择

就是这样一种进化机制，博伊德在演讲中提到了这一点。就他所做的论述，我会做一个评论。这是其演讲中篇幅相对较短的一部分内容，但能极好地说明文化进化理论与其他进化模型在推导方面的区别。例如，在基因系统的模型中，群体选择并不被重视，因为迁移行为会对群体基因的一致性造成一定的破坏。而且，文化进化理论不仅激发了一般公众的想象，还引发了一系列学术讨论。

文化群体选择的假设检验

文化群体选择这个推论很难得到验证，不仅因为其本身在很多时候所指代的事物是不同的，还因为群体结构既可以是基因层面上的，也可以是文化层面上的，还可以兼而有之。所谓群体，可能是说同一种语言的一群人，也可能是住在同一个村镇里的一群人，甚至是一个团伙；成员之间的关系是建立在基因或文化基础上的。人类的行为多种多样，很多都不能作为衡量群体特性的指标，但是人类学的核心是研究不同文化群体之间的差异，如果不同文化群体在行为上没有明显差异，那么人类学也就没有存在的意义了。我确实认为，在某些行为和制度的进化方面，文化群体选择发挥着非常重要的作用，而且群体间的竞争理论上是一种重要的选择力量。因此我认为，现在主要问题出在实证方面。我们需要做的是，找出哪一种行为在群体层面上发生了变化，以及发生了多大的变化；群体间的竞争会选择哪一种行为特性；是如何选择的；是否有足够的时间让群体产生显著的文化变化。

通过文化的群体选择，战争会促成群体内的利他主义。这种假设似乎已经成了一种广为人知的理论，在科学杂志甚至流行杂志上都能看到。这种观点倒是与我们的直觉很吻合。父母常跟我谈起“敦刻尔克精神”。第二次世界大战时，很多英国民众冒着巨大的风险，驾驶自家的小船去营救被困在法国海滩上的英国士兵，因为英国军队被入侵法国的德国军队击败了。我父亲那时候还是个小男孩，住在萨福克郡（Suffolk）的海边，认识不少当年驾船实施跨海营救的人。当时，我嫂子的父亲是应征入伍的士兵，在海峡彼岸排着长队等待救援。虽然敌人的炮火尚未停息，但士兵们接到的命令依然是，如果不好好排队就会被上级开枪击毙。如此看来，英国军队为了维持纪律，对“惩罚”和“利他”给予了同样的重视。关于这一点，在所有和敦刻尔克有关的电影中，我都没有看到过。因为有了这场营救，丘吉尔得以将一场失败的战役转化成了一次重要的宣传活动，在一个令人绝望的时刻传递出了一种精神，而这种精神

甚至改变了战争的进程。

我们可以提出一种替代性的假设，即个体利益是人们选择参与战争的最重要的因素。图尔卡纳部落的勇士们在参与掠夺牲畜时肯定会考虑个体利益。未参与掠夺的人自然不能参与战利品的分配，而没有牲畜就没办法结婚和繁衍后代，因此一些年轻人会从战略角度出发，对这一问题进行思考，并很有可能得出“这是一种值得冒险的活动”的结论。所有军队都会对那些拒不参战的人进行惩罚，以杜绝临战脱逃之类的行为；宗教或其他机构的规范也都会涉及对违规者的惩罚。虽然惩罚是由群体选择进化而来的一种规范，但一些领导者会利用不正确的权力来强制执行；在这种情形下，惩罚便与群体选择或利他主义的强化脱离了关系。可以看出，惩罚是群体选择的结果，但不同群体的惩罚是否不尽相同、不同群体是否会因为成功程度的不同而进化出惩罚措施，并使其成为符合公共利益的规范，我们目前还不得而知，或许将来也很难知道。从自然设计角度出发进行推导，往往是检验进化假设的首选方式，但不见得是最强有力的方式。

我感兴趣的是该如何检验这样一种观念：群体间的竞争和冲突，比如种族间的战争，会导致狭隘的利他主义，也就是说利他对象仅限于群体内部成员。为了验证这一假设，我的博士生安东尼奥·席尔瓦（Antonio Silva）对贝尔法斯特（Belfast）地区持续了300年之久的天主教共和党（Catholic Republican）与基督教统一党（Protestant Unionist）之争进行了研究。1998年，《耶稣受难日协议》（*Good Friday Agreement*）达成之后，冲突大体上归于平静，但并没有完全消停下来。在西方许多地方，政府修建的隔离墙依然耸立，不同宗教之间依然鲜少通婚，冲突事件依然存在。当地的许多街区都画有大幅的宗教壁画，教徒们的游行队伍也会高举宗教画报，一面鼓舞士气，一面挑衅对方。教育也是分开实行的，要么是天主教学校，要么是基督教学校。当然，以旁观者的角度来看，这些都有待做进一步研究。

在具体的研究过程中，我们采用了试验法。那些愿意填写调查问卷的人会得到5英镑的酬劳，并可以将这笔钱捐给当地的学校，要么是天主教学校，要么是基督教学校，或者中立的慈善机构。调查内容包括基本情况、宗教信仰程度，以及是否经历过宗教主义威胁等（Silva & Mace，2014）。

我们在贝尔法斯特市的21个地区进行了试验，有的地区冲突情况较严重，有的地区则较平和。我们还委托这些参与者把一些“遗失的信

件”送到其所在的天主教或基督教机构。预计，由于宗教主义威胁的增加，与控制组，即中立的慈善机构相比，被试组外部的利他主义行为会减少，而内部的利他行为会增加。在冲突地区，无论是被试组还是控制组，对外部群体，也就是非自身所属群体的利他行为，如归还信件、向学校捐赠等都会降低，这与预测的一致。战争会使人憎恨敌人，但是冲突并没有增加群体内部的利他行为。被试组与控制组的内部利他主义基本上处于同一水平状态。

田野工作完成后没多久，贝尔法斯特地区就爆发了一场严重的骚乱，起因是双方就贝尔法斯特市市政大厅里的每一种旗帜每年应该悬挂多少天发生了争执。我们立刻折返回去，在上次做过调查的一些地区重新进行试验。结果显示，在冲突期间，各种类型的利他行为都有所下降（Silva & Mace，2015）。虽然冲突是导致利他主义产生的因素之一，但所起到的作用是很小的。

这一结果着实令人惊讶，因为根据其他一些学者的研究，因冲突而在经济领域展开对抗的群体，内部的利他行为会有所增加。通过仔细研究，我们发现，那些学者所提供的证据不够清晰。一项研究表明，2006年，以色列养老院中的老人们在参加博弈试验时，对在经济上不够慷慨的行为做出惩罚的概率远超战前和战后；但在这项试验中，既没有控制组，也没有外部参照组（Gneezy & Fessler，2012）。

另外一项研究的对象是格鲁吉亚的学生。在对待同学方面，那些有过较多战争经历的学生会比那些来自“远方的学校”，没怎么经历过战争的学生表现得更为慷慨（Bauer et al.，2014）。但是，在这项研究中，有过战争经历的成人与没有经历过战争的成人在行为方面并没有明显不同。总之，很少有研究像我们的研究这样，针对真正意义上的群体，关注真实冲突中向真实机构进行捐赠的问题，对冲突双方进行检验，而且还设置了控制组。

大量研究已证实，合作行为不仅会高度依赖具体环境，还会随着群体的变化而变化，而且对当地的生态系统也极为敏感。我并不能确定，合作行为是由文化群体选择进化而来的众多行为之一。随着时间的推移，与合作行为相比，其他一些行为或习惯在同一种文化内或者不同文化间表现出了更为稳定的多样性。

正如博伊德所观察到的那样，系统进化法，譬如建立在相似性基础上的语系进化法，是测量文化规范持久性的有效方法之一。已知的很多文化特性，比如婚姻和亲缘关系等方面的规范和准则，以及政治体制，

都非常强烈地表现出了文化史系统的进化迹象，这表明这种机制确实对文化特性在不同文化群体中的变异起到了强化作用。

因此，对于很多文化特性来说，习惯更容易成为群体的选择。另外，个体依据使用频率来进行行为选择，也是一种可以强化文化变异的机制，可参考中国亲缘关系系统的进化（Ji et al.，2016）。不同的行为具有不同的进化路径，希望通过越来越多的明确的实证研究，这些具体的路径能被揭示。我们当前所接触到的只是一个大型实证项目的表层，随着项目的推进，需要开展的工作还有很多。

06
适应性、合作性、操控性和敌对

保罗·西布莱特

法国图卢兹大学经济学教授

罗伯特·博伊德在演讲中非常坚定且诙谐地揭示出人类是这个星球上最具生态环境适应性、合作规模最大的物种。当然，人类社会中的竞争，就残暴性而言也是所有物种中最突出的，操控同类的狡猾本领也是所有物种中最厉害的。这样的描绘看似前后矛盾，但事实上，后两种特性是前两种特性的必然产物。真相是，超级合作性不仅能使人类创造出巨大的资源和收益，还带来了竞争和操控。

如果人类的生活是一场零和博弈，那么人与人之间的互动将会趋于简单直接，人们会避免做任何对他人有益之事，这样一来，互动程度将会降至最低。然而，人类生活实则是一场正和博弈。人们经常，而且理应去做有益于他人之事，因为只有这样，才能对合作所获收益做出预期。因此，在对合作收益进行分配时，如果分配结果未达到预期，我们就很容易陷入被剥夺状态，然后，通常会进行报复，就像俗语所说，“君子报仇，十年不晚”。对于那些曾经是朋友，现在却变成敌人的人们来说，应该最明白个中滋味。

无疑，在各种人类社会群体中，个体的合作对象会不断发生变化，这是合作关系的核心所在。这一事实进一步强化了上述几种特性之间的张力。合作很少会发生在群体层面上，而常发生于亚群体和各种联合体内部，因此，群体的构成很容易发生变化。在这种情况下，个体一方面要与自身所在的亚群体或联合体中的部分成员进行合作，另一方面还要面临竞争，并需要通过竞争让自己脱颖而出，以便被群体中更具实力的联合体接纳。所以我们常会遇到这样的情况：与部分个体合作的同时，还要与另一部分个体竞争；与某个个体在某种条件下合作，同时又在另一种条件下竞争。合作与竞争之间的张力会导致非常复杂的动机和行为。对于人类的这一阴暗面，博伊德并没有否认，但也没有展开讨论。在本篇短评中，我想解释一下，为什么我认为这部分被他忽视的内容其实很重要。

博伊德的系列演讲是对他和彼得·里克森，以及一些同事和学生，比如约瑟夫·亨里奇和萨拉·马修，共同开展的一个重要研究项目的完美总结。[1]社会科学和人文领域的研究者应该对这些研究成果给予更多的关注，因为它们深刻地揭示了人类的本性，对人类社会的运转机制有着强大的启示。该项目旨在揭示为什么人类成了自然界中超乎寻常的存在，

为什么人类能够在地球上任何一个有生命的地方存活。该研究结果认为，这要归功于人类的超级合作能力，是它使人类在群体层面上更为聪明，而这种聪明的程度远超任何一个个体。事实上，个体并不需要像群体这般聪明，因为我们的群体进化出了最有效的信息传播机制，比如模仿，而且在进行模仿时，并不需要理解所模仿的行为模式背后准确的因果关系。

按照博伊德的说法，这些特性不仅能帮助人类找到解决生存问题的技术性方案，还能帮助人类掌控环境。此外，还能以集体行为规范的形式帮助人类解决集体行动问题，即公共产品生产。这些集体行为规范会通过一些社会机制得以强化，比如对违规者的惩罚。这些规范会随着群体间竞争不断进化，在这一过程中，那些更具适应性的规范会更受群体选择的青睐，从而得以在群体层面上传播开来。博伊德及其合作者们将这一过程称为“文化群体选择”。关于这一称呼，最大的问题在于，它与生物学上的群体选择过程毫无关系。[2]

地图是对现实世界的简化，以方便人们了解世界。与此类似，所有的模型都需要对所需建构的现实进行简化。虽然地图有时会误导使用者，比如遗漏了某个重要的道路交叉口，但我们不能因为它对现实世界进行了简化而批评它。博伊德将人类所面临的生态挑战和社会挑战简化为了如何激励个体参与公共产品生产。由此可见，如果能够生产一些公共产品，那么群体将能得到更好的发展。同时，如果群体领袖能通过一些过程做出一些决策，那么群体将会衍生出具有强制性的规范，来对个体在公共产品生产过程中应做的贡献做出规定，并必然会对个体进行监督。如果在监督过程中发现有个体没有做出应做的贡献，也必然会对其进行惩罚。人类学家萨拉·马修对图尔卡纳人所做的研究表明，在生产具体的公共产品时，这一过程确实有用，即使这种公共产品的生产是“参与对敌对群体的武装突袭”。

接下来，我所关注的是与社会生活有关的具有误导性的论述，虽然这种误导性只存在于部分领域之中。来看一些被遗漏的关键特征，主要体现在以下三个方面：

1. 一般来说，监督个体是否遵守规范要比公共产品一例所显示的要困难得多，这就意味着，群体成员往往很难对“谁违规”“谁没有违规”达成一致意见。

2. 规范不仅需要强化，还需要沟通。群体成员通常很难对适用于

一般情况或特殊情况的规范达成一致意见。

3. 虽然强化某些规范比没有规范更能给所有群体成员带来公共利益，但是群体成员往往很难做出抉择，到底要在哪个潜在规范上达成一致意见，甚至会因此而产生巨大的利益冲突。这就意味着，一些个体会为了自身利益而对信息的传递过程进行操控，从而影响规范的强化过程，也就增加了强化规范的难度。

接下来，我将分别针对这三个方面进行论述。

通过直接制裁来监督对规范的遵守情况

正如博伊德所观察到的那样，“借助大规模群体内部的对等性来强化合作”这一观点存在很大问题。这种对等性指的是，群体内部的个体面临着这样一种威胁：如果有人不提供公共产品服务，那么其他人也会效仿，拒不提供公共产品服务。这种观点的问题在于，在确认谁是违规者时经常会出错，一旦出错就会导致更多违规行为出现。

博伊德认为，这一问题的解决之道是直接制裁，而且是很精准的直接制裁。他写道，在这种情况下，“失误或不确定性并不会激发一系列的背叛行为，这就意味着，在一个喧嚣的世界中，合作是可以稳定存在的”。显然，博伊德已经有很多年没有研究诸如“找出哪些孩子会给家庭或学校带来麻烦”之类的问题了。直接制裁可能会致使人们把心思花在争论“谁是违规者”“谁不是违规者”这些事情上；而规范的监管者很难直接判断出谁违了规，谁没有违规。

博伊德在好几处都做出了暗示，在大规模群体内部存在着一种群体智慧，能够认定违规者，但没有清晰地表达出这一观点。实际上，这是完全不可能的。基本上，所有围绕着“谁违背了哪种规范”而展开的争论，注定会和肥皂剧一样不了了之。在人类生活的这一核心领域中，人们是会撒谎的；越是热切地想要强化某种规范，围绕事实所展开的争论就越难理出头绪。

规范沟通和规范强化的重要性

那些寻求强化社会规范的人需要确认所有群体成员都对规范有所了解。这一点很重要，其原因有二。第一个原因是，虽然规范看似是普适

性的，且存在于不同类型的人类社会中，所有社会都有杀戮、保卫、性与婚姻，以及食物制作等方面的行为规范，但不同社会中的规范的具体内容却千差万别。在两个相邻社会中，食物禁忌很可能大为不同，婚姻制度也可能不一样。相邻群体间的不同规范很多时候也是族群的重要标志，例如，食物禁忌和安息日等规范时刻提醒着犹太人，自己与当地的其他人属于不同的族群。重要的沟通仪式有时候会通过很精妙的方式告知和提醒群体成员，他们需要观察并掌握何种规范，然而很多时候这都是徒劳的，因为按照上一条原则，这些规范没有那么“显而易见”。因此，沟通越精妙，就越有可能产生错误，人们也越有可能执行错误的规范。

规范沟通之所以重要的第二个原因是，即便一些规范是普适性的，但这些规范在具体运用中也会发生变化。例如这样一个规范：为了保护家庭成员和同胞，男人们应该挺身而出参与战斗，就算失去生命也在所不惜。据我所知，几乎在所有的社会中都有这种规范。然而，需要在何时挺身而出，谁需要做出牺牲，做出多大的牺牲等一系列问题是个体无法自我决断的，因此大多数社会都建立了某种权威机构，以进行某些类型的沟通，比如告知群体成员如何运用某些规范。就军队而言，基本规范通常主要是由军队领导人给出的；当然还存在一些深化的规范，主要是关于何时、何地、如何遵循领导人指示的。这些规范往往充满了争议。这又一次表明，在对规范的强化过程进行简化时，不能做出这样的假设：在一个群体中，群体领袖能轻易地察觉出谁违了规，并对其进行恰当的惩罚。

有人认为，只有规模庞大且复杂的现代社会才会孕育出与“什么是合适的规范”有关的争论。这种说法虽然很吸引人，但有证据表明，这类争论出现的时间其实非常早，甚至可以追溯到文字尚未出现的时代。最近，系统发生学分析发现，一个名叫《史密斯与魔鬼》（*The Smith and the Devil*）的民间传说可以追溯到青铜器时代（Graca da Silva & Tehrani 2016）。这个民间传说最有意思的一点是，它将两种强大的道德规范置于对立面，第一种规范是讲真话，第二种规范是拒绝与坏人合作。这个传说的主题是如何在一个充满矛盾的规范世界中行事。该主题也是荷马史诗《伊利亚特》和《奥德赛》的核心主题，亦是印度两大梵语史诗《摩诃婆罗多》（*Mahabharata*）和《罗摩衍那》（*Ramayana*）的核心主题。毫无疑问，这样的争论并不是现代社会的产物。

最重要的一点是，个体在很多时候都不清楚，自己在何些特殊情况

下需要遵守何种规范。这就为那些别有用心的个体提供了通规范来操控他人的可能性。

规范强化过程中的利益冲突

“边缘之外”（Beyond the Fringe）喜剧团队创作了这样一个滑稽剧：第二次世界大战时期，彼得·库克（Peter Cook）所扮演的英国皇家空军指挥官劝说乔纳森·米勒（Jonathan Miller）所扮演的狂热的爱国者珀金斯（Perkins）执行自杀式攻击：“珀金斯，我们现在应该豁出去。这会改变这场战争的基调。”珀金斯接受了这项任务，并在登上飞机前说：“长官，再见！还是应该说，回头见？”毫无疑问，他得到的回答是“不，我们不会再见了，珀金斯”。

这些意味着，在大多数社会中，大部分围绕规范所进行的沟通都会涉及某些个体对另一些个体的操控，但博伊德所描述的澳大利亚扬德鲁万塔人在制作渔网和处理大柄苹时所进行的规范沟通，并未体现出这种操控性。在这些例子中，群体成员因为拥有共同的利益，所以会诚实地进行信息传递。很难看到，一位上了年纪的扬德鲁万塔人通过扭曲处理大柄苹的信息来获得收益。

但是，因为拥有共同利益而诚实地传递信息，这种情况并不是人类社会的常态。人类学家如果把这种非常态视为常态，就会低估自然选择所赋予的认知机制的复杂性。如果认为史前时期的进化对人类的要求很简单，也就是说人类只需盲目模仿所在的狩猎群体中的长者所告之的需要模仿的事情即可，那么自然会推断出，这正是今人在面对操控者时容易上当受骗的原因所在。如果认为史前时期已存在大量这样的操控者，而我们的祖先也已学会如何和这类人打交道，那么就需要重新审视一下，为何当下的人们依然容易被操控者欺骗。

为了证明上述观点，我将对神话和民间传说进行分析，以期揭示出：对于群体间的信息传递而言，那些试图以隐瞒真相来获益的群体会对信息传递过程进行操控，而不会进行忠实的传递。已有许多学者对神话和传说中常见的情节进行过分析，比如著名作家约瑟夫·坎贝尔（Joseph Campbell，1949）[11]和克里斯托夫·布克尔（Christopher Booker，2006）。他们通过各自深刻的文学洞察力，对许多神话和后神话的情节进行了分析，并发现它们拥有共同的结构。

然而，他们对这种共同的结构所做出的心理学解释是完全不合情理

的，因为他们基本上采用的都是荣格的原型概念（Jungian Archetypes）[12]。不过我们可以用另一种更令人信服的假设来对这种共同的结构做出解释，从而说明为何这些情节会如此雷同，而且这种假设与博伊德及其同事们所进行的研究具有一致性。博伊德所提供的心理学解释要比荣格的好：情节所涉及的背景可以无限延伸，但情节所涉及的内容必须有所选择。虽然这些情节在结构上大体相似，但真正重要的是细节。之所以这么说，是因为在这些情节中，细节被赋予的任务不是强化“应该遵循某种规范”之类的一般信息，而是要帮助受众应对规范冲突所带来的挑战，从而让人们明白在何种情况下应该遵守何种规范。换句话说，这些情节反映出：在很多情况下，个体并不清楚自己应该遵循何种规范，而这些情节帮助他或她（在大部分神话中是他）明白该怎么做。

在这里列举几例。不难发现，所有这些情节所传达出的信息都不支持这样一种观点，即情节只是对社会中占主导地位的规范的强化。

1. “打败怪物”之类的故事情节。无论是约瑟夫·坎贝尔，还是克里斯托夫·布克尔，都认为这是最基本的故事情节，通常会被描述为英雄找到了某种制胜怪物的方法，而非依赖超自然力量。这类情节所传递的信息并不是“人们应该坚守，不能退缩”，如果是为了强调“为了保卫群体而勇敢战斗”这种规范的话，那么人们会很容易联想到上述信息。实际上，这类情节是在说，人们需要通过努力思考找到击败强大对手的最佳方法。

2. 许多戏剧和文学作品都会设定一个厉害的权威人物，但其地位早晚会被英雄人物取代。

3. 在那些关于“禁令”，比如不要打开房门，不要吃某种水果之类的故事中，总会出现“英雄对禁令的无视”之类的情节。这虽会带来危险，却不会导致灾难。如果是为了强化社会规范，那么无视禁令必然会导致灾难。实际上，这些和“禁令”有关的故事让人们明白了该如何对待禁令，以及如何破除它。

4. 直到近期，应该说直到19世纪之前，戏剧和文学故事中几乎都没有出现过“女性人物进行欺骗”之类的情节。欺骗是人物施展权力的重要特征之一，但鲜有女性会这么做，除了希腊神话中的皇后和女神。显然，如果是为了强化社会规范，就不应该强调社会权威人物所具有的欺骗性。

总之，情节并不只是用来强化某种社会规范的，更重要的是为个体提供思考机会：在各种困境中该如何运用规范。情节中往往会对相关规范进行清晰的描述，以此与其他冲突性规范进行对比，或与该规范的其他运用方式进行比较。规范沟通要求沟通者都具有强大的社会认知能力，因为沟通过程是很容易受到操控的。

在此，我需要总结一下上述围绕规范沟通所做的讨论，以便回应博伊德在演讲中提到的“文化沟通”观点。文化传播是策略性行为，在传播过程中，信息的传递者会试图操控信息接收者，虽然并不总能获得成功。有时候，传递者与接收者是拥有共同利益的，比如大部分狩猎技术的传授者与接收者。然而，狩猎技术的传承并不是文化传播的典型实例。与狩猎技术的传承完全不同，围绕规范所展开的争论涉及更多的利益冲突，沟通者很少有利益一致的时候。在人类的文化史上，这类现象可以追溯到相当久远的时期。事实上，如果认为文化操控是现代社会独有的现象，那么将无法理解文化战争为何会如此令人头痛。

如果把这种沟通看作一种复杂的、具有操控性的活动，那么就能更好地理解，为何在可以借助集体智慧的情况下，个体智力依然会变得如此之高。个体只需拥有普通的智力便能在自然环境中行动自如，因为文化会帮助人们完成大部分工作。但是，要想在社会环境中如鱼得水，就不能只依赖与他人的合作，还需要掌握更多的技能，以使自己变得足够聪明。

A DIFFERENT KIND OF ANIMAL

下 篇

回应篇

07

文化、观念和决策

谢谢四位评论人，他们的评论是经过深思熟虑的，所提出的观点也很具挑战性。他们对我的文章提出了一些严肃且认真的批评，但并没有说服我，至于为什么，接下来我将做出说明。不过，我依然很感谢他们能如此严肃认真地参与这样一个基础的研究项目。这个研究项目的核心是运用种群生物学的理论和模型来研究文化及文化进化。里克森和我刚开始进行这项研究时，氛围可不是这样的；尽管受到了很多批评，但大部分批评都与我们的核心论点毫无关系。

很多批评都是这样的："这不是我的学科对文化的思考方式。"还有些批评不太礼貌。20世纪80年代早期，一位非常知名的人类学家应邀在一个大型会议上对我的一篇文章进行评论，他说："我多希望当初没有把时间浪费在这篇文章上。"从那以后，批评之声越来越多。当然，比较幸运的是，并不是所有人都持批评态度。另外，利用种群生物学理论对文化进化进行模型建构，也不再让人觉得太具革命性。虽然四位评论人对我的部分观点感到很不安，但至少认可探究文化进化的价值。无论是谁，对文化进化的看法大多都有可能是错误的，但实实在在的讨论肯定能让我们对文化进化有更深入的理解。

四位评论人中的三位，科学哲学家金·斯特林、经济学家保罗·西布莱特和人类学家鲁思·梅斯都指出：我在研究人类行为时，对人类"搜集信息并做出明智决策的能力"举证不足。就这一点而言，他们的观点代表了一大批学者对人类社会的看法。如果对社会科学领域的学者们的主成分进行一次分析，毫无疑问，第一种或第二种主成分关注的肯定是：个人决策在社会发展过程中所扮演的角色。在这个轴的左边，将会是以社会学家爱弥尔·涂尔干（Émile Durkheim）为代表的大多数20世纪文化人类学研究者，他们强调的是社会如何创造个体；在轴的右边，将会是社会学家托马斯·霍布斯（Thomas Hubbes）、现代理性选择学派，以及行为生态学家们，他们将社会视为个体选择的产物。斯特林、西布莱特和梅斯都在轴的右边。斯特林认为，人类是非常聪明的，知道自己应该从他人那里学习些什么；西布莱特认为，人们可以巧妙地操控行为规范；梅斯则认为，诸如图尔卡纳部落的突袭行动之类的合作行为可以用个人利益来解释。

对这些"位于右边"的思想家的诸多观点，我都是赞同的。人们通过

努力工作来获得更多个人利益，任何忽视这一事实的理论都会铸成大错。问题在于，人类在做决策时通常会依据自身对世界的看法，而“右边的理论”往往无法解释为什么人们对世界的看法会是这样的。这一点在经济学的主流模型，即现代理性选择模型中表现得尤为突出。行动者的个人偏好被抽象为效用函数（utility functious）[13]，而信仰则被抽象为贝叶斯式的先验假设。在处于平衡状态，且他人选择既定的情况下，行动者会选择能够将自身预期效益最大化的行为。

经济学家通常会假设人们拥有正确的先验条件等，并认为大部分人都知道自己正在做什么，也知道如何做出合理的选择。其他学科虽然没有表达得这么直白，但基本上也都如此认为。

我认为，这里需要注意的是，人们是从自己的生活中获得对世界的认知的。我在年轻时相信，吃完饭一小时后才能游泳，否则就会抽筋，甚至淹死在水里。我之所以相信这个观点是因为身边的每个人都相信。我现在相信，吃完饭马上就能游泳，因为医学专家说这么做没问题，但是新观念可能会给我的孩子们带来极大的风险，甚至死亡。其实这两种观念都缺乏直接证据。我的确看到过有孩子一吃完饭就去游泳，然而，如果这么做会有1%的死亡风险，我该怎么办？我无法接受这种风险，而以自身经验又无法完全排除它。我母亲坚信，知道一个人的星座就能了解一个人的性格。她之所以相信这一观点是因为她周围的人都相信。她还认为，她的观察证实了这个观点的正确性。我不认同这个观点，因为我周围的人都不认同。

我和学生们花了大量的时间研究进化与行为之间的关系：200年前，人们探讨的可能是“三位一体”的教义问题。毫无疑问，自我观念一旦确立，个体选择理论就会变得非常实用。但是这样一来，在回溯历史时，就需要解释人们为什么会确立这样的信仰，而解释又需要依赖特定的社会背景。在我看来，我在文章中所阐述的方法足以解释，人类为何以此种方式被塑造。

回应艾伦·奥尔

艾伦·奥尔在其评论中提出了两个重要问题：文化学习与认知能力之间的关系；我和彼得·里克森的研究工作与奥地利经济学家哈耶克的学术观点之间的关系。在文章中，我确实没有提及文化学习与智力之间的复杂关系，也没有对我们的研究工作与其他学者的工作之间的关系进行说

明。因此，非常感谢奥尔提供了这样一个机会，让我可以就这两个问题展开论述。

我赞同奥尔的观点，也就是说，有充分的理由相信，认知能力和文化学习是互相强化的。强大的认知能力能提高文化学习的效率，而文化学习则能通过积累来提升认知能力。文化学习在本质上属于推理的范畴。认知能力的提升，尤其是心智能力和因果推理能力的提升，有助于人们对他人行为或事件关系做出更准确的判断。与他人有关的心智能力被称为心理理论，在传授知识时尤为重要。

试想一下，一位初学者应该如何向榜样学习？榜样会有一些心理表征，比如认为“鸟”可以指代知更鸟，也可以指代鸽子和鸵鸟等，但不能用来指代蝙蝠。在观察榜样行为的过程中，当榜样说“看木槿树上的那只鸟”时，学习者必须推断出榜样头脑中想的是什么。[1]人类的心智比其他灵长类动物进化得更完善，因果推理、物理学常识以及和物质有关的直觉等，不仅能够帮助一部分人创造出有用的工具，同时还能帮助其他人复制这些工具。

而且很有可能，文化学习会促使人类进行认知能力方面的投资，从而带来更大的适应性。向他人学习的能力帮助文化学习者接触到比个体学习多得多的知识，进而接触到更多的数据，提升认知能力，并使这种能力更具价值。通过文化所获得的知识表征，也有助于人们去学习很多原来不感兴趣的知识。例如，在指代数字的单词出现后，人们便能追踪大于“3”或“4”的整数了，[2]而这种能力开启了多种新的认知可能性。

这一推论意味着，在认知能力和文化学习之间存在着正反馈关系。遗憾的是，化石和考古学资料所提供的，关于人类祖先认知能力和文化学习本质属性方面的证据十分少。大约在200万年前，南方古猿被人属物种所取代。这些人属物种个体的大脑容量比南方古猿的大脑容量大50%，[3]因此他们能够制造更为复杂的石器，而且足迹遍布欧亚大陆。但是，其发展速度与猿的发展速度基本一致，这意味着文化学习在其生活中并不重要。在接下来的150万年间，人属物种个体的大脑容量变得更大了，但从发掘出的石器来看，其生活中并没有出现累积性的文化进化。当然，考古记录总是不完整的，因为它无法保留那些容易消失的技术，比如修建居所和船只、去除食物毒素、制衣与烹饪等。30万年前的石器已经比较“现代”了，已体现出和某些技术有关的零星线索，这表明当时人类的生活已在慢慢接近现代生活。

如果要打赌的话，我敢肯定，人属物种的出现与文化学习能力的进

化在轨迹上是一致的；在接下来的100万年中，文化学习能力和认知能力彼此契合，相互提升，并最终达到了一个临界点，使文化积累得以成为可能。由于考古记录普遍存在缺失，对文化学习背后的认知机制的研究也存在太多不确定性，因此在打赌的时候不用考虑钱财的得失，对此我非常开心。

现在我要谈一下，以往学者对我们研究项目的贡献。我们在加利福尼亚大学戴维斯分校教授“环境研究导论”的时候，这个研究项目诞生了。我们原计划开设一门主题为“个体的目标追求会导致环境的恶化，且任何人都难辞其咎”的课程。我们知道，在生物学和经济学中有类似的观点。彼得·克里森是生态学家乔治·帕金斯·马什（George Perkins Marsh）的支持者，因此想研究前工业社会中的污染问题和滥用资源问题。然而，当时在浏览主流的生态人类学文献时，我们却感到不太满意，或许是因为年少气盛，觉得自己能够做得更好。文化进化理论当时还没有成型，所以我们阅读了进化生物学家泰德·克罗克（Ted Cloak）和人类学家尤金·鲁伊尔（Eugene Ruyle）的不少文章，以及进化生物学家罗恩·普里亚姆（Ron Pulliam）和克里斯·邓福德（Chris Dunford）合著的一本杰出作品。心理学家唐纳德·坎贝尔（Donald T. Campbell）所写的一篇和文化进化有关的论文极大地影响了我们。后来，我们利用进化生态学方法，尤其是生态学家理查德·莱文斯（Richard Levins）所提出的方法做了一些创造性工作。种群遗传学家卢卡·卡瓦利-斯福尔萨和免疫学家马克·费尔德曼的论文让我们领略了模型在群体遗传学和流行病医学领域内所发挥的力量。卡瓦利-斯福尔萨和费尔德曼非常大方，向我们开放了在斯坦福大学开设的一门课程，该课程的主要内容来自他们在1981年出版的一本书。加利福尼亚大学戴维斯分校的种群遗传学家迈克尔·图雷利（Michael Turelli）和生物学家约翰·吉莱斯皮（John Gillespie）所教授的群体遗传学课程也让我们受益良多。此外，我还有幸能在数量遗传学复兴时期前往芝加哥大学生物学家麦克·韦德的实验室学习了一年。我们的研究方式过去是，现在依然是广泛阅读人类学、心理学和经济学等领域的文献资料，挖掘有前瞻性的经验问题，而后用种群生物学理论来解决它们。

我们也很关注社会科学领域中的一些非常重要的先驱者。艾伦·奥尔一针见血地指出，我们的工作与奥地利学派的经济学家哈耶克的工作存在契合之处。但在仔细阅读过哈耶克的著作之后，我发现他的基本研究思路与我们有很大不同，所以我认为最好把更多的时间花在自己的研究项目上。类似的情况还出现在19世纪晚期的社会学家加布里埃尔·塔尔德

（Gabriel Tarde）身上。很有可能，我们的研究结果及观点与哈耶克、塔尔德以及其他一些先驱者的观点如出一辙。若真如此，那就让未来的学者来判定，荣誉应该归谁吧。

回应金·斯特林

我有一条生活法则：永远不要与哲学家辩论。这条法则源自我与杜克大学的生物哲学家鲍勃·布兰登（Bob Brandon）就“选择发生的层面”所进行的辩论。20世纪80年代早期，我对选择理论的了解与常人无异，因此每一次辩论都以失败而告终。接二连三的打击让我深刻地明白了：千万不要与专业人士辩论。然而，对于金·斯特林的批评，大部分我都不认同。接下来我就解释一下，为什么我认为他错了。

有一点是斯特林和我都认同的，那就是群体智慧总需要依赖个体智力。毫无疑问，即便是“穿过拥挤的城市街道”之类的简单任务所要求的计算能力也超出了大部分复杂的机器学习算法能企及的范围。当然，在这一点上，蝴蝶、狒狒和人类没什么不同。

“独特的人类智慧”究竟为何物？斯特林的说法有点含糊。是否存在一些方法可以有效地测量出，将人类与其他物种区分开的一般认知能力？答案颇具争议。但我认为，有足够的证据表明，人类在有关工具制造能力方面的因果推理和心理理论等特定任务上的表现，比其他哺乳类动物好。[4]因此很有可能，正是这些方面的智慧造就了人类在进化上的成功。

斯特林可能认为，我的观点是“只靠文化就能解释人类在进化上的成功”。黑猩猩如果能进行精准的社会学习，那么很快就能制作出梭镖投射器和勒瓦娄哇尖状石器（Levallois point），但我的观点其实并非如此。人类拥有许多关键的显性性状（phenotype），比如认知能力、双足行走、双手灵活、童年期较长以及青春期发育迅速等。对于人类来说，所有这些都是必不可少的，而且是不够的。文化适应也是生物表现型的基本要素之一。显然，斯特林对此并无异议，所以我在这里就不做赘述了。

斯特林声称，文化传播“在很大程度上属于简单的依照惯例的启发式策略”。我所理解的其所说的“依照惯例”是指，学习者掌握某种特性的概率与其所观察到的这一特性的使用人数之间存在着一种递增关系[5]，而其他一些因素，比如个体经验和进化而来的心理机制等，则对掌握特性

的概率[6]影响不大。如果这是他想要表达的，那么他肯定错了，我对此并不认同。

第一，正如斯特林所看到的那样，我认为社会学习者会更关注模仿对象的特质，也就是所谓的“模仿成功者”。这一点在文化学习中是至关重要的，意味着不能把文化学习视为“简单的依照惯例的启发式策略”。“模仿成功者”的倾向会推动学习者以不均衡的方式学习“成功者的特质”，由此带来的结果是，这些适应性特质将得以被传播，即便学习者并不清楚它们是如何造就成功的。对方言、创新的扩散等方面所做的研究，以及各项实验均表明，文化学习者使用的是一种基于成功的启发式策略。[7]此外，有证据表明，有些学习机制会增加学习者加入某些群体的概率，例如，群体成员与学习者相熟，或是群体成员展示出了更多的可靠性，等等。各种机制共同造就了文化学习，因而学习者更有可能获得适应性特质。你不能骑墙，如果认为“基于成功的学习倾向”更重要的话，就不能认为文化传播只是简单的依照惯例的启发式策略。

第二，人类认知机制的诸多方面共同造就了学习方式及其倾向性，从而使人们更容易学习某些事物。我在本文中并没有花太多时间讨论这一论题，但曾和彼得·里克森在别的文章里着重探讨过。[8]认知机制的某些方面的确会使文化学习带有倾向性，即使人们倾向于学习那些能够增强适应性的行为。例如，人类学家克拉克·巴雷特（Clark　Barrett）曾通过研究指出，相较于“哪些动物可食用”之类的信息，儿童更容易记住“哪些动物很危险”之类的信息，[9]这可能是因为他们会觉得知道“什么是危险的”更为重要。心理学家安妮·沃茨（Annie　Wertz）也发现，儿童不太喜欢吃蔬菜，这可能是因为有很多植物与大柄苹一样有毒。[10]

我在文章中所描述的数学模型并不存在此种“内容”[11]上的偏差，虽然我假设两种环境具有同等的可能性和均衡的适应性，但这仅仅只是假设。如果忽略这两个被简化的不太贴合实际的前提条件，该模型给出的预测结果会是：学习将会具有一定的倾向性，也就是说，学习者会倾向于掌握更具适应性的特质。[12]需要注意的是，这种倾向性会将文化进化引导至更具适应性的方向上，但就其本身而言，并不足以生成可造就人类成功的复杂的文化适应性。即便人类倾向于学习“何种植物有毒”，但这种倾向性并不会告诉人们要避开大柄苹，以及还需要掌握地方性知识。

另外，在大部分的文化学习中，特定的学习倾向背后的认知机制往往进化自其他一些因素。文字系统只出现了几千年，而在几百年前，大

部分人还都是文盲。因此，大脑神经系统中掌管阅读的部分在被自然选择造就之初，并非是基于认读需求的。正因如此，神经系统才会对文字系统造成诸多限制。[13]

金·斯特林的另一个错误在于，他认为我的观点是：作为行为主体，鲜有人能够以较低成本找到可靠的直接环境信息样本。在我所讨论的模型中，环境信息的准确程度是一个可变参数。研究证明，当环境信息准确时，受自然选择青睐的学习规则通常会忽视社会信息。因此，对于那些获得了充分的环境反馈信息的行为而言，模型的预测结果是，社会信息的作用是次要的。这与斯特林的观点一致。模型假设，在很难学习到最佳行为时，人们将以模仿他人为主要学习方式。毫无疑问，生活充满了各种情况，而个体会在各种情况下从环境中获得大量反馈，并据此调整自身行为。对于人类的适应性来说，这是非常重要的。

路易斯·利本伯格在论述“动物踪迹追踪”时提供了许多绝佳的例子。此外，还有一点也是毋庸置疑的，那就是许多技能需要通过持续练习来完善。斐济人惯用一种长弯刀来完成许多事务，包括建造房屋。无论是砍削出柱子和房梁，还是削平茅草屋顶，这些技能的掌握都离不开大量的练习。可以肯定，只依靠观察别人怎么做，是远远不够的。确实存在许多一学就会的特质，但也有许多特质是需要大量练习才能掌握的，这样的例子不胜枚举。在这一点上，斯特林是对的，但超出了讨论范围。

我的观点是：人类要想在各种杂乱的，或是缺乏反馈信息的环境中获得成功，就必须重视各种特质；而自然选择塑造了人类的心理机制，因此在学习某些特质时，人们会选择加入正在展示这些特质的群体中。在试图掌握这些特质时，人们会将特质的共性作为线索，也就是斯特林所谓的“依照惯例的启发式策略”，以及对内容进行练习，对榜样进行选择。能够获得充分的环境反馈信息的例子，并不能证明我的观点有错。要想证明我是错的，就得先证明缺乏环境反馈信息的特质很少，或是证明这类特质对人类的适应性无关紧要，或是证明人类在掌握这些特质时没有使用任何社会信息。

如果人们更常依赖社会信息，那么将出现许多会导致错误适应的文化特质，斯特林对这种看法表示怀疑。如果人类不具有模仿倾向，那么这种看法将是正确的。当然，这并不是我的观点。“模仿成功者”之类的规则将摒除那些会导致错误适应的观念；某种观念越容易导致错误适应，就越有可能被摒除。个中道理很微妙，那是因为，与自然选择一样，向榜样学习的倾向性在强弱程度上取决于群体中的变异量。当大部

分人都在做相同的事情时，向榜样学习的倾向性就会很微弱。这意味着，一种错误的适应行为若由于某些原因而变得非常普遍，那么就可能会持续一段时间。因此，我建构的模型会做出这样的预测：对于观察到可导致错误适应的特质，除非它们被广泛传播，否则我们不用为此而付出太高成本。

和非适应性文化传播特质有关的例子也有很多。就连斯特林也承认，人们受到了许多不正确的宗教信念的影响。不过在他看来，这些错误信念并不会给人们造成太大损失。然而，我却认为献身此类宗教会导致独身主义、自我否定以及其他一些高代偿行为。正如斯特林所注意到的那样，各类民间医疗体系似乎都含有许多无用甚至有害的用药方式与疗法。在许多地方，疾病和厄运常会被归因于亲朋邻里的恶意行径，而此类信念又常会引发一些代价极大的行为。例如，很久以前，居住在巴布亚新几内亚弗莱河流域的一支吉卜赛人认为，死亡皆由巫术所致。一旦有人死亡，他们就会进行调查，利用一些支离破碎的线索指认某人犯了重罪，并将之处死。[14]纵然拥有较多财富，但在工业国家中，大部分人仍保持着小家庭的生活方式，并不会把全部财富都花在享乐上。[15]若是一一列出，那么非适应性行为的清单或许会很长。[16]

还有一种可能是，很多行为虽然有用，但并非最优。它们未经严格检验却得以延续至今，因为是通过社会学习的方式获取到的。人们依赖复杂的人造制品，尽管它们会带来复杂的设计难题。斐济人的房屋拥有许多特性，建造者在材料、结构和尺寸等方面做出了正确的决策。正如斯特林所注意到的那样，茅草厚度之类的特性是需要参考反馈信息的。下雨时，屋顶是否会漏水？不过这么说也不够准确，因为茅草的成本和可获得性会随时间的推移而变化，当成本较高时，“屋顶稍有漏雨”便成为最优选择。台风来袭时，房屋的结构至关重要。若是多年没有台风，一旦台风突至，恐怕没有哪栋房屋能幸免于难，当然也不排除会碰上好运。毕竟，这些房屋的构造并不是很复杂。

就算是在小规模社会中，人们也会依赖一些复杂的工具，比如独木舟、复合弓、回力标等。人们在使用许多技术时，或许并不清楚其中的因果关系。树薯，也被称为木薯，是一种高产的根茎类食用作物，被广泛种植于南美洲的热带地区。某些树薯品种含有大量的氰化物，为了去除毒素，南美土著采用了各种各样的技术。[17]这些技术一般都有多个步骤，需要持续数天，还需要大量的劳动力；而且，他们并不能说清楚每个步骤对于去除毒素而言究竟有何效用。

那么，我们从何可知，这些复杂性会促使人们采取最优行为？有个证据是：居住环境类似但文化背景不同的群体所制造出的人工制品也不同。传统的汤加房屋在设计上与传统的斐济房屋大为不同。当然，不排除存在一些细微的环境差异使汤加房屋在汤加地区更适用，斐济房屋在斐济地区更适用，但这种可能性很小。斐济房屋的基础设计被广泛应用于斐济群岛的各个岛屿，而就这些岛屿而言，无论是台风到访频率，还是森林覆盖率以及其他一些会影响房屋设计的因素都不尽相同。这就意味着，这种房屋设计依据的并不是当地环境的具体情况。斐济群岛东部的一些岛屿在20世纪的大部分时间里都属于汤加海洋帝国，所以在这些岛屿上，斐济人建造的全是汤加房屋。由此可见，房屋建造技术会与某些文化因素相结合，而后为当地人所掌握。

其他的文化特质也一样。欧洲人来到南美洲之后，将当地的树薯带到了非洲，但处理树薯的技术却未能在非洲应运而生。这意味着，仅仅利用一些线索是很难学会技术的。[18]更普遍的例子是，利用语言特征来预测技术和生活实践，要比利用环境特征来预测来得更准确，这也说明，技术和生活实践并不会在每一代人身上都得到优化。

关于树薯的例子表明，对文化线索的依赖是一种强大的机制，能够增强人类适应环境的能力。由于处理树薯的技术非常复杂，而且其中的因果关系不甚清晰，所以个体学习起来非常困难，再加上环境反馈极其有限，未经处理的食物所导致的后果有时候需要过一段时间才会得以显现。如果打算先了解实践活动是如何带来益处的，而后再采纳行为，那么将无法掌握“有效去除毒素”之类的方法。另外，社会线索有助于群体逐渐进化出一些能让生活状态愈加丰富的实践活动。

斯特林还错误地理解了水獭、白蚁和变形虫等动物所构建的复杂事物与本人观点的相关性。毫无疑问，那些都与学习无关。关键之处正如心理学家艾莉森·高普尼克（Alison Gopnik）[14]等一众认知心理学家所认为的那样，对因果关系的理解是必不可少的，如果不具备此种能力，那么所构建之物就无法得到改进，也就无法适应当地环境的变化。其他动物所构建的事物证明，精准的基因遗传和无声的自然选择过程将会促生一种算法，使其他动物在构建高度设计化的事物时，能做出调整以适应当地的具体情况，而不需要具备任何洞察力。无疑，精准的文化传播和无声的“模仿成功者”的过程会带来同样的效果。

回应鲁思·梅斯

鲁思·梅斯和我曾在观念上有过一些小冲突，但对于她所谈到的大部分内容，我都比较认同，对此我很开心，尤其是她对进化社会科学领域的发展及现状所做的观察，令我倍感欣慰。各学科内部的自我中心主义是人文科学发展的强大敌人。本科生在选修经济学、人类学和心理学课程时，会接触到许多和人类行为、人类社会有关的观点，而这些观点有时甚至是水火不容的。理解人类本非易事，所以出现上述情况情有可原。

但是，令人无法原谅的是，竟然有那么多学者坦然接受了人文科学的现状。多年来，进化学家们重新树立起了一个微缩格局，以人类行为生态学、进化心理学和文化进化理论三个学派为主，然而这几家几乎一直在自说自话。据我所知，我们中的许多人都对此深感不满。我们应该在共同的进化观基础上，去解决不同学派间因差异而导致的问题。对此，同梅斯一样，我也认为在过去几年间已经出现了转机。梅斯就是这一综合研究取向的主要代表人物之一，运用系统遗传学方法研究人类行为的变异。许多研究者，尤其是年轻一代的研究者，可以轻松地从心理学领域跨入有关适应性的群体动力学领域。

接下来，我将探讨文化进化理论和人类行为生态学的关联，所涉及的观点都源自梅斯对我文章的评论。累积性文化适应改变了适应性的产生规模。人类行为生态学通常依据的事实是，文化适应性使不同的人类种群得以进化出不同的局部适应性。正如梅斯所说，人类行为生态学家试图解释各人类群体间的行为变异。他们所提出的大多数假设都基于这样一种模型：人类能让基因适应度在所处的局部环境中达到最优。

人类行为生态学家偏好使用“最优适应度”这一观点，而不在意这些变异是否是由进化而来的基因差异和文化适应性，以及个体学习所致。这一情况主要源于这样一种观点：自然选择通常会形成一种发展进程，进而使结果达到最优适应度，无论其对环境中的偶然性是否敏感。人类行为生态学家很清楚，有时候近似机制起到了很重要的作用，但不要忘了，最优适应度是一种强大的启发式策略。另外，我认为生物学的经验能证明这些观点是正确的。我同意梅斯的看法，那就是累积性文化适应通常会让人类的适应度达到最优。

人类行为生态学家有时候会把研究工作定义为：适用于人类一般行为的生态学。事实上这是不准确的。行为生态学家在研究其他物种时所关注的是不同物种间的差异，而非同一物种的不同种群间的变异。当然，这并非是说，不同种群间不存在行为差异。如我们所知，非洲西部

一些种群的黑猩猩会使用石头敲开硬壳的坚果，而另一些种群的黑猩猩则不会。不过相对于人类种群间的变异来说，这种差异的规模相当有限。我们急需对人类种群间所存在的巨大行为变异做出进一步解释。

我认为，人类种群间之所以存在如此巨大的变异，是因为人类具有文化适应性。如果不具有文化适应性，人类就和其他物种并无二致。学习、推理，以及其他形式的个体适应性，使得种群间产生出一定程度的差异，同时人类认知能力的局限性又影响着这种局部适应性，而正是这种局部适应性最终将各人类种群区别开来。有些人类行为生态学家，不包括梅斯，对文化在变异产生过程中所起的重要作用有质疑。譬如一位著名的人类行为生态学家在阅读完《文化与进化过程》一书后对我说："对这么点发展给予了过多的关注。"然而，假如我说得没错的话，人类行为生态学家所研究的物种间变异是有赖于文化适应性的。

梅斯的研究完美地阐释了上述观点。人类行为生态学家尝试着从生态学角度去探究促生人类种群间差异的因果机制。例如，牧区经济是否会促使人们选择父系继承制度？为了回答这一类问题，梅斯搜集了大量和人类种群有关的资料，以研究在牧区经济和父系继承制度之间是否存在统计学上的相关性。人类学家早已明白，要做到这一点，就不能将拥有共同文化史的不同群体视为独立的取值点，只是尚未找到系统的方法来克服这个难题。比较生物学领域也存在类似问题，而生物学家们找到了一些复杂的统计学方法来处理这些难题。梅斯借鉴这些方法成功地解释了文化变异的相关问题。该研究工作隐含的假设是，人类种群的历史发展轨迹类似于非人类种群的进化史。人类的文化适应性如果只依赖学习或个体知识的累积，就无法通过任何系统惯性得到矫正，而系统发生学的方法也就失去了存在的意义。

接下来要讨论的是梅斯提出的那个具有挑战性的问题：如何对文化群体选择这一假设进行检验。毫无疑问，我十分赞成梅斯的看法，文化群体选择理论已经得到了很好的阐释。此外，我认为有许多可靠的实例可以支撑文化群体选择理论。当然，我也承认，更系统的实证检验方法的确必不可少，不过前提是这些检验方法需要被正确使用。在这一点上，我的观点与梅斯的意见是有分歧的。

梅斯认为，图尔卡纳人参与战斗的主要原因是个体利益需求，而非文化的群体选择。然而，这算不上是替代性假设。文化群体选择模型表明，群体内部的行为会受到个体利益的驱动。奖惩措施能促使人们遵守规范，并使规范稳定下去。如果遵守规范无法使个体受益，那么文化群

体选择假说就会失效。一个参加突袭行动的图尔卡纳人在战斗打响时缩头缩脑，待到分配战利品时，如果缺乏强制性的监督机制，人们就很难阻止他拿走一定份额的战利品，而这并不符合任何人的个体利益需求。在相互监督的情况下，一般来说，遵守规范有利于全体成员的个体利益。个体利益可以用来解释为何人们会遵守规范和执行规范，但并不能用来解释规范的形成。这就好比，图尔卡纳人会谴责突袭本部落同胞的行径，但同时对其他部落的突袭行动却有增无减。文化群体选择假说认为，规范的形成可以用群体间竞争来解释。当群体处于竞争状态下时，个体利益需求可以使规范稳定下来。

换句话说，文化群体选择模型并未显示出，狭隘的利他主义会得到进化。[19]梅斯对自己所做的实证研究进行了描述，并表示，至少在实验中，群体内部冲突不会促生利他主义行为。这项研究非常有趣，因为与另一些研究者所观察到的结果正好相反。梅斯的研究结果并不支持这一假设：人们具有经由遗传进化而来的倾向性，因此在遇到外部攻击时，群体成员会表现出更多的合作行为。但是，梅斯统计出的数据和文化群体选择理论无关，因为文化群体选择理论推断，由自利性动机所维持的规范也会给群体带来益处，且不需要人们变成利他主义者；文化群体选择模型还推断，真正重要的群体边界是文化边界。

在梅斯看来，在检验文化群体选择模型时，不能仅限于验证由模型设计所得出的推论。我同意她的这一看法。目前，这类实证研究的确存在，但并不十分确切。[20]但我并不认为，单靠这类研究的结果就能否定文化群体选择理论。科学研究理应呈现出百家争鸣之态。如果没有其他竞争性推论的存在，我们也就无法对“文化群体选择促生了规范”这一假设进行验证。在这一方面，人类行为生态学可能就心有余而不足了。据我所知，人类行为生态学家至今尚未提出过任何令人信服的推论来解释规范的形成。

梅斯认为，领袖的权威性能够解释人们何以能拥有当下的规范，但这一看法显然是不对的。即使在最简单的觅食社会中，规范都有赖于相当数量的成员的共同遵守。在诸如图尔卡纳部落这种食物生产型社会中，规范是由上万人甚至数十万人共同遵守的。在这种大规模群体中，只有形成足够的权威性才能使个体接纳并执行新规范，而个体和小团体都实难做到。

此外，在漫长的人类社会史中，非政府控制的社会状态一直占据着主导地位。在这种社会状态下，文化变异的规模往往都很大，以致任何

政治机构都不可能就规范的内容做出决策。经济学家在个体利益最优化的基础上构建出了和规范内容有关的进化模型，但是该模型并不能对群体利益产生帮助，而且进化过程缓慢得令人难以置信。由于缺乏某些关于群体间竞争的潜在推论，因此我们应该沿用文化群体选择模型。

回应保罗·西布莱特

保罗·西布莱特认为，我围绕规范所做的阐释忽略了三个特征，而这些特征恰好表明，规范并没有我所说的那样有效。这三个特征分别是：第一，违规行为的判定；第二，强制性规范的判定，人们对此存在争议；第三，利益冲突会导致规范在执行过程中受到人为操纵，且对操纵者有利。西布莱特的观点无疑是对的。这些特征的确都是削弱合作稳定性的因素，但我不确认，他是否认为这些特征都会破坏合作。一方面，他描绘了一个非常黑暗的画面，让人以为非正式的强制性规范永远都不会起作用；另一方面，他又承认非正式的强制性规范的确能够解释图尔卡纳人之间的合作。

那么，我就来阐述一下个人想法。如果群体规模较小，且具有相同的文化特质，那么西布莱特所提出的问题就不难克服。在具有多种文化特质的大规模群体中，非正式的强制性规范并不能起到稳定合作的作用。复杂社会中的大规模合作一般需要经过上万年的进化才能形成规范。这些规范都非常新颖，而且更加复杂。不过，在小规模的觅食部落或者小农庄里，非正式的强制性规范则是可以促成合作的。接下来就解释一下，为何规范的监督和沟通，以及利益冲突在文化同质的小规模社会中并不会引发大问题；然后再探讨一下，我与西布莱特在对待文化传播问题时所表现出的一些重要的不同之处，以及为何说西布莱特高估了那些阴暗面的重要性。

先来讨论一下有关监督的问题。我曾经常常需要从我的几个孩子中找出真正的“肇事者”，不过那已是多年前的事情了，就这个方面而言，西布莱特的看法是对的，因为他的孩子与我最小的孩子差不多大，不过我不能确定他是否算得上是一个教育专家。当然，这并不是问题的关键。成人总想将自己认定的规范强加到孩子身上，但我们并没有生活在儿童世界里，也不知道那里到底发生了什么事情。

同成人世界一样，孩子们的世界也充满了规范和调控。在我儿子六岁的时候，我们从郊外搬回了洛杉矶。在此前差不多一年的时间里，他

都生活在那个偏远的地方，只跟其他两个孩子有来往。在回到洛杉矶几周后，一天，上一年级的儿子放学回家，看起来闷闷不乐。我问他发生了什么事情，他叹了口气说："规则太多了。"通过进一步的交流，我终于明白，他所说的"规则"并非来自老师，因为老师制定的规则其实很简单；而是指在校园中发挥作用的各类不成文的规定，而这些规定让他很是苦恼。孩子们会相互监督并执行规范，有时候甚至会让人觉得冷酷无情。

同现代化的、匿名的、都市化的、多民族的社会生活相比，乡村生活更像校园生活。那个世界不大，但生活通常却很开放，人们拥有更多了解外界的机会，而且对此兴趣浓烈。我曾对斐济群岛上的一个村庄做过研究，在那里，当你遇到一个人时，打招呼的标准方式是："你要去哪儿？"作为一个城里人，我对此的第一反应是："不关你的事。"但是对于村里人来说，这种回应是非常粗鲁无礼的，因为在那里，你去哪儿、做什么事情，关系到每个人的生活。在小规模群体中，所发生之事皆会被众人悉数了解到，而且还会被口口相传下去。比如谁逃脱了惩罚，哪里又发生什么秘密事件，谁与谁有染，哪里又被盗了，等等。对于某些人的所作所为，人们经常会发生争执，最后再通过规范机制来平息争议。尽管如此，人们依然深知，做了坏事就会被抓。当群体规模足够小且人们相互熟识时，群体智慧能发挥出最大效用。

为了了解规范的运行机制，不妨来看一下，觅食部落日常生活中的那些促进合作的规范。分配食物，尤其是分配肉类，在狩猎-采集社会的经济活动中发挥着至关重要的作用，可以说是一种保障措施，能有效降低大型狩猎活动的风险。在许多部落中，有关规范会明确指出，哪些人可以分得猎物身上哪些部位的肉。例如，前文所提到的刚果盆地的姆班吉拉部落中，杀死猎物的狩猎者可以分得猎物的心脏，其他狩猎者可以分得肾脏，等等。在南部非洲卡拉哈里（Kalahari）沙漠中的一些部落里，杀死猎物的狩猎者只能分得肋骨和一块肩胛骨，年轻男性只能分得肾脏、腹壁和生殖器部位。还有些部落不允许杀死猎物的狩猎者来主导分配。例如，在卡拉哈里地区的朱·霍安西（Ju/'hoansi）部落中，是谁制造了杀死猎物的弓箭，就由谁来分配猎物，而狩猎者无法主导分配。[21]一位老人在沙漠营地里屠宰一只黑羚羊，而身边围满了着急想得到自己份额的人；若是有人拿到的肉超出了其应得份额，那他肯定会被揭发。

我想你应该不会觉得这很不可思议。你或许会想，如果那些人躲在

灌木丛中偷食猎物，就不会有人知道他们作弊了。这种情况确有发生，但是根据我在亚利桑那州立大学的同事金·希尔的描述，在阿齐部落中，此类违规行为仍然常常被揭发出来，因为阿齐人个个都是森林中的神探，可以根据犯罪现场留下的痕迹推断出是谁在此做了什么事。非常有趣的是，无论是在刚果还是在卡拉哈里，人们都会接受这样一种文化观念：违规者在今后的打猎活动中会很倒霉。这种真伪难辨的观念会激励人们按照规范行事，从而使规范得到巩固。

西布莱特指出，涉及公共利益的合作是不难被解释的，因为人们拥有共同的利益。我则认为，公共产品生产是最难被解释的，因为在大规模群体中，互惠准则无法发挥作用。不过，要清楚的是，分配肉食不但和公共利益无关，而且还会导致参与者之间发生激烈的利益冲突。作为较高水平的狩猎者，年轻男性如果仗着自己身强力壮独吞了猎物，那么就会被众人唾弃。分享行为的确提供了长期利益，虽然你可能认为这种规范对于维持小规模合作来说并非必须。对于狩猎-采集部落的食物分配来说，关系亲密度和互惠性都发挥着重要作用，不过相关证据表明，除了这些因素之外，规范也同样发挥着重要作用。[22]一些灵长类动物很少分享食物，当一只黑猩猩捕获到猎物时，其他黑猩猩就会在它身边相互推搡或乞哀告怜。

西布莱特认为，规范是需要人们达成共识的，这非常重要。我同意他的这一看法。如果人们在是非问题上意见不一，麻烦便会随之而来。正如他所说，现代中东地区的冲突即是一个令人沮丧的例证。但这并不意味着，就具体的规范进行有意识的沟通有多么重要。实际上，在“面对面”的社会中，就规范进行的有意识的沟通，在规范的传播过程中只能起到一定程度的作用。在西布莱特看来，“文化传播是策略性行为，在传播过程中，信息的传递者会试图操控信息接收者，虽然他们并不总能获得成功”。有关文化传播的此类描述在经济学中很常见，[23]但与我从进化论取向的角度，对社会学习所做出的阐释并不一致。更重要的是，这种描述与关于“文化的实际传播过程”的实验研究结果也不一致。

最重要的是，文化学习是由信息接收者主导的，而非信息发送者。儿童并不是空空如也的容器，任由成人灌输文化信息。他们拥有与生俱来的心智，能够从社会环境中提取有用知识。在大多数社会中，幼童会向父母及其他家庭成员学习，因为那些人是随时都能接触到的。稍大一点的孩子们会在家庭之外度过大部分时间，会观察村子里正在发生着什么，还会学习周围人的言行，例如语言学习，通过倾听而掌握上千个词

汇的含义。知识的传授顾然重要，但这并不意味着要说服他们使用“dog”而不是“bund”。

儿童的注意力和成人并无二致，他们会在注意力的帮助下缩小假设条件的范围，而后依据这些条件去思考词汇的含义。许多适应性技能都是这样被逐渐掌握的。人们做着自己愿意做的事，儿童通过观察学习着如何做这些事，进而参与其中。[24]一个小女孩观察到，所有编织席子的人都会在特定的时间收集露树叶，并会采用特定的技术来编织。讲授主要是用来帮助学习者了解“什么是重要的”，以及掌握一些比较复杂的手工技能，比如编织。儿童不会任由成人操控，因为他们会观察日常生活中人们的言行，并会用成人的实际行为来检验其所说的话。儿童在观察到孕妇忌食大型的掠食性鱼类时，因为知道放弃这种富有营养的美味很难，所以会相信这些妇女所认为的“这些鱼类对孕妇有害”。

如果接受口口相传的观念，就要承担被操纵的风险。不过，坚持言行一致这一原则可以使风险最小化。当有人告诉你，不遵守规范就会在以后的狩猎活动中走霉运时，如果他是一位从未参与过狩猎的老人，你可能会怀疑他所说的话，但如果他是一位颇有成就且一直恪守规范的狩猎者，那么这种口头教导的风险就小多了。无论是实验室研究还是田野工作，结果均表明，学习者倾向于采纳言行一致者的行为。[25]

这就意味着，在文化同质的社会中，人们往往不用在不同规范间进行选择，只需采纳在社会中占主导地位的规范。斐济人认为，身体位置高于他人是很不礼貌的，而其许多社交活动都需要席地而坐，因此在进出房间时人们都得采用爬行的方式。儿童也会经历这种场面，并从中学习到相关规范；对于他们来说，这些规范不是成人告之的，而是通过自身观察获得的。斐济人还有些非常强硬的规范，比如不允许异性表亲之间有任何社交活动，包括讲话。[26]孩子们会从成人的实际行动、偶尔的疏忽和违规，以及羞耻感中观察到许多规范。事实上，大部分规范都不是通过沟通和传播被强加给孩子们的，而是被他们自己采纳的。

规范在时空维度上的演变也支持了这一观点。上述斐济人的规范在斐济群岛的各个村庄里随处可见。欧洲人初来乍到之时，斐济群岛上的总人口大约为30万。个体决策显然无法促生适用范围如此广泛的规范。西布莱特提出，这些规范是通过某种政治过程而为大众所接受的。但是，在小规模社会中，被人们共同遵守的规范在传播范围上，通常会远大于社会中任一政治机构所制定的任一规范。在采集-狩猎社会中，通常不会存在比团队更高级别的永久性机构。在复杂社会中，比如在成为殖

民地之前的斐济群岛上，政治组织的规模约为几千人。波利尼西亚（Polynesian）社会进入分裂状态已有上千年，并且已扩张到东太平洋地区；它也拥有共同规范，例如建立在等级森严的父系氏族制度基础之上的政治体制。

除了时空维度这个因素，其他某些因素也会促使一些技术或环境发生突变，进而引发规范的突变。在公元六七世纪时，马匹被引入大平原地区，从而引发了该地区印第安部落的巨变。以觅食为生的科曼奇人、以种植玉米为生的克罗人（Crow），以及北方森林的觅食者夏延人（Cheyenne）等都陆续迁移到了大平原地区，并通过融合建立起了新的社会系统。这个社会系统基于“男人社会”规则，拥有大型的夏季营地、频繁的战争以及夏季启动仪式。

我同意西布莱特所认为的，不同规范之间的冲突由来已久。规范并不像“不可杀人”这样简单。大量人种志证据表明，规范亦具有偶然性。[28]当时是否属于自卫？是否喝醉了？是否和受害者是近亲？受害者是否做了一些有辱家门之事？规范实则非常复杂，且拥有一系列微妙的算法，这些算法可以用来处理偶然事件。[29]尽管如此，规范之间时而也会出现冲突，但只要冲突不常发生，规范便可以正常地运作下去。

与西布莱特不同，我不太喜欢用民间故事来解析以往的社会。根本原因在于，这些民间故事的内容并非源自某种社会功能，而是源自某些事件的趣味性或纪念价值。[30]有充分的理由怀疑，民间故事的内容因此而出现了偏差，所以才常常涉及冲突。这就好比在未来的某个时刻，某人试图根据作家阿加莎·克里斯蒂（Agatha Christie）的小说集重构20世纪的英国生活；他或许能从中找到一些准确的信息，但肯定会高估谋杀案出现的概率。

我和彼得·里克森的合作长达40年，这些演讲即是合作的产物。对我来说，里克森亦师亦友，他博学多识，拥有丰富的社会学和生物学知识。我认识的人无出其右。同时，他还是一位极富洞察力和创造力的思想家，而且为人和善、注重实践、慷慨大方。我很难找到比他更好的人。

我以前的学生也让我受益匪浅，尤其是约瑟夫·亨里奇。他沿袭了里克森和我所提出的一些观点，并把它们变成了巧妙的实证研究项目，还提出了一些很好的想法。我很高兴，我这个年龄的人无须达到亨里奇所设定的标准。如果你想学习更多有关文化进化的知识，他的书《人类成功统治地球的秘密》（*The Secret of Our Suuess*）是一个极好的选择。

在与其他一些极其优秀的学生的合作中，我也学到了很多。他们是米歇尔·克兰、理查德·迈克埃尔尔里斯（Richard McElreath）、克里斯蒂娜·莫亚以及卡西克·潘钱纳森（Karthik Panchanathan）。

最近有幸能与同事萨拉·马修合作，他是一个非常睿智的思想家，也是一位极具天赋的、高产的田野工作者。第2章中的很多观点都源自他的研究结果。

5年前，我离开加利福尼亚州大学洛杉矶分校，来到了亚利桑那州立大学，得以与人类起源研究所的研究者们结缘。这真是一桩天大的幸事。我们获得了外界的资助，开展了一个跨学科的研究项目，旨在探究人类何以如此独特。在项目进行过程中，我从人类起源研究所的同事们那里，尤其是金·希尔那里获益良多。他是在世的研究狩猎-采集社会的最伟大的民族志学家，不仅思维缜密，而且极富创造力。本书中的很多观点都来自与希尔的讨论。柯蒂斯·马里安（Curtis Marean）拥有早期现代人类考古领域的渊博学识，比尔·金贝尔（Bill Kimbel）在研究更新世原始人方面拥有百科全书式的知识，两位都给了我很大帮助。与人类起源研究所的其他同事，如伊恩·吉尔比（Ian Gilby）、凯文·兰格格雷伯（Kevin Langergraber）、查尔斯·佩罗、凯·里德（Kaye Read）、加里·施瓦茨（Gary Schwartz）、安妮·斯通（Anne Stone）等人的交流也让我收获颇丰。我的博士后，马西克·胡德克（Maciek Chudek）、麦克斯·德雷（Max Derex）和希拉里·伦菲斯特（Hillary Lenfesty）是我很多灵

感的提供者，他们帮助我保持着学术上的警醒。

很多人都看过本书的初稿，并提出了非常有用的意见和建议，包括心理学家克拉克·巴雷托（Clark Barrett）、人类学家约瑟夫·亨里奇、人类学家摩西·霍夫曼（Moshe Hoffman）、心理学家吉利恩·乔丹（Jillian Jordan）、普林斯顿大学出版社执行编辑艾莉森·卡莱特（Alison Kalett）、伦敦大学教授鲁思·梅斯、普林斯顿大学人类价值研究中心主任斯蒂芬·马塞多、人类学家克里斯蒂娜·莫亚、进化生物学家艾伦·奥尔、进化生物学家大卫·兰德（David Rand）、经济学家保罗·西布莱特、科学哲学家金·斯特林，谢谢你们。感谢我长期的登山伙伴、优秀的散文家约翰·威利（John Wiley），他所做的编辑工作极大地提升了本书的文字水平。

需要感谢的人还有很多，尤其要感谢我的妻子琼·西尔克（Joan Silk）。35年来，她一直在家里举办“灵长类动物社会进化”研讨会。她很善于将我那些不成熟的想法清晰地表达出来，经过她的编辑，那些枯燥无味的表述总能变得赏心悦目，本书亦复如是。

01 人类的独特之处（一）：模仿学习

[1] 严格来讲，这种果实是植物的孢子体，也就是产生孢子的种子状结构。

[2] Howitt et al., 1862; Cathcart, 2014; Moorehead, 1963.

[3] Nowak, 1999, 664.

[4] Aplin et al., 2003.

[5] 物种生物量即在给定时间内属于该物种的所有个体的数量之和。

[6] Hill et al., 2009.

[7] Binford, 2001.

[8] Kay & Hoekstra, 2008.

[9] Collias & Collias, 1964.

[10] E.g., Klein, 2009.

[11] Wadley et al., 2009.

[12] 同上。

[13] Tennie et al., 2009; Dean et al., 2014.

[14] Collias & Collias, 1973.

[15] Tebbich et al., 2001.

[16] 与欧洲探险者的迷失经历有关的实例还有很多。在博伊德等人于2011年出版的一本著作中，富兰克林探险队（Franklin expedition）的经历跃然纸上；在约瑟夫·亨里奇于2015年出版的著作中，也有一些相关实例。

[17] Howitt et al., 1862; Cathcart, 2014; Moorehead, 1963.

[18] Cathcart, 2014, Kindle location 2158-2160.

[19] Earl & McCleary, 1994.

[20] Clarke 2012, Kindle location 1029.

[21] 同上。

[22] Wadley et al., 2009, 9593.

[23] Tooby & DeVore 1987; Barrett et al.; 2007 & Pinker 2010.

[24] 社会科学领域的研究者们普遍认为，人们之所以采取某种行动，是因为他们明白该行动为何比其他行动更好。在经济学模型中，行动者投入成本并完成创新，其他行动者则会采纳这些创新，因为他们明白这些创新的工作原理，并深知其好处。对此，经济学家肯尼斯·约瑟夫·阿罗（Kenneth J.Arrow）于1962年、经济学家罗伊·鲁道夫·罗默（Roy Rudolf Romer）于1993年所提到的和宗教法规有关的事例可作为参考。心理学家和哲学家大多也持类似观点，同时这也是常见的社会观念之一。

[25] Dawkins, 1986.

[26] Henrich & Henrich, 2010.

[27] Broesch, 2010.

[28] Boyd & Richerson 1987, 1995, 1996; Perreault et al., 2012.

[29] Perreault et al., 2012.

[30] 当总体样本数量足够大时，抽样所致的总体样本结构的随机变化便可忽略不计。

[31] 有两个平均值需要追踪。第一个是个体间任一时刻的平均值，这个平均值一直在发生着变化，因为环境每隔T代就会变化一次；每次变化之后，大部分个体随即会表现出某些非适应性行为，此时，较大的g值就会受到青睐；而后，随着表现出适应性行为的个体在比例上逐步占优，较小的g值便会受到青睐。第二个是在一段相当长的时期内，经过多次环境变化后的平均值。这个平均值是一个相对稳定的常数值，并会长期受到自然选择的青睐。

[32] Denham et al., 2009.

[33] E.g., Boyd & Richerson 1985, 1996.

[34] Basalla, 1988; Petroski, 1992a.

[35] S. Johnson, 2010.

[36] 与此有关的更为全面的论述请参阅约瑟夫·亨里奇和弗朗西斯科·吉尔-怀特（Francisco Gil-White）发表于2001年的论著，以及亨里奇发表于2005年的著作。

[37] 更多事例与实验数据请参阅彼得·里克森和罗伯特·博伊德发表于2005年的著作，以及约瑟夫·亨里奇发表于2015年的著作。

[38] Henrich, 2009.

[39] Henrich, 2015.

[40] Hansell, 2005.

[41] Collias & Collias, 1964.

[42] Gould & Gould, 2007; Hansell, 2005.

[43] Jacklyn, 1992.

[44] Gould & Gould, 2012.

[45] Enquist et al., 2008.

[46] Perreault, 2012.

[47] Gingrich, 1983.

[48] 请参阅进化生物学家尼克·莱恩发表于2009年的论著。对于横纹肌而言，或许是两次。另请参阅社会学家乔治·斯坦梅茨（George Steinmetz）等人发表于2012年的著作。

[49] Lane 2009.

[50] Tennie et al., 2009.

[51] Boyd & Richerson, 1996.

[52] Whiten et al., 2005.

[53] Hill et al., 2014.

[54] Gergely & Csibra 2006.

[55] T. Morgan et al., 2015.

[56] Boyd & Mathew, 2015.

[57] Boyd & Richerson, 1996.

[58] Mathew & Perreault, 2015.

[59] 萨拉·马修和查尔斯·佩罗通过主成分分析法适当减少了预测变量的维数，以使不同类别中的预测变量数据具有可比性，并以此总结出了不同类别中预测变量的相同变化量。他们利用模型对统计的变量进行了比较，并针对每个特质确定了相应的最优逻辑回归模型，然后将文化历史、生态系统和距离的标准化β值相加，生成了每个特质的总体度量。

[60] Jordan, 2015.

[61] Currie et al., 2010.

[62] Holden & Mace, 2003.

[63] Shennan, 2001.

[64] Henrich, 2004, 2006.

[65] Powell et al., 2009.

[66] Kline & Boyd, 2008.

[67] Derex & Boyd, 2015.

[68] Cosmides & Tooby, 2006.

[69] Richerson & Boyd, 2005, Henrich, 2015.

[70] Beppu & Grififths, 2009.

[71] Boyer, 2001.

[72] Henrich & Gil-White, 2001; Henrich, 2015.

[73] Henrich & Gil-White, 2001; Boyd & Richerson, 1985.

02 人类的独特之处（二）：社会规范

[1] 在网上可以观看。

[2] Petroski, 1992b.

[3] Seabright, 2010.

[4] Greenhouse, 2014.

[5] Kaplan et al., 2000.

[6] Hill, 2002.

[7] 同上。

[8] Wood & Gilby, forthcoming.

[9] Lukas & Clutton-Brock 2012a, 2012b.

[10] 在网上可以观看。

[11] 更新世大约开始于260万年前，结束于11 700年前。

[12] B. Smith, 2011.

[13] Bliege-Bird et al., 2009.

[14] B. Smith, 2011.

[15] Clastres, 1972.

[16] Steward, 1933.

[17] Wilke, 2013.

[18] Brink, 2008.

[19] Brink, 2005; Friesen, 2013.

[20] Brink, 2013; O'Shea et al., 2013.

[21] Swezey & Heizer, 1977.

[22] Gat, 2015.

[23] E.g., Fry, 2006; Manson & Wrangham, 1991.

[24] Gat, 2015; LeBlanc, 2014.

[25] J. Morgan, 1852.

[26] Gat, 2015.

[27] Sutton, 2014.

[28] M. Smith, 1938; Hämäläinen, 2008.

[29] Bamforth, 1994.

[30] Darwent & Darwent, 2014.

[31] James & Graziani, 1992.

[32] Kroeber, 1976.

[33] Darwent & Darwent, 2014.

[34] Couzin & Franks, 2003; Franks, 1986.

[35] E.g., Sterelny, 2012; Pinker, 2010.

[36] Iwaniuk & Arnold, 2004.

[37] 只有当自然选择的力量非常微弱，且会对其他某些看似合理的条件加以控制时，r、b和c的定义才是准确的。详细介绍请参阅科学哲学家乔纳森·伯奇（Jonathan Birch）和萨米尔·奥卡沙（Samir Okasha）发表于2014年的论著。

[38] 要明白为什么，可以假设，参与建造一个地洞或驱赶线之类的公共设施，个体会损失1个单位的适应性，但群体成员将人均收获0.1个单位的适应性。再假设群体成员为100人，个体广义适应性的净变化为$0.1 \times 1 + 0.1 \times 99 \times r - 1$，其中$r$是个体与其他群体成员之间的相关系数的平均值。若与群体成员没有关系，则$r = 0$，那么个体广义适应性的净变化值即为$0.1 \times 1 + 0.1 \times 99 \times 0 - 1 = -0.9$，此时自然选择并不支持该个体参与合作。若与群体成员为兄弟姐妹，那么个体广义适应性的净变化值则为$0.1 \times 99 \times 0.5 + 1 \times 0.1 - 1 \approx 4$，此时自然选择会支持个体参与合作；当然，这种情况常出现在昆虫之类的动物群体内，鲜见于哺乳类动物的群体中。

[39] Boomsma et al., 2011.

[40] Hill et al., 2011.

[41] Hill et al., 2014.

[42] 同上。

[43] 亲缘关系取决于群体的规模和迁移率。规模过大或迁移率过高，都会削弱亲缘关系。究

其原因，不妨设想一下，在一个群体中，随机选取两名群体成员并追溯其家族史，然后可能会看到以下两种情况：第一种情况是，他们拥有某个共同的祖先，也就是说，这两个人在遗传上一脉相承；另一种情况是，其中一人拥有移民血统，因为群体人口众多，所以难免会有来自远方的移民。可以肯定，移民与群体成员没有共同的祖先。亲缘关系可以简单地理解为，这两个人具有共同祖先的概率。群体规模越大，所需要追溯的年代就越久远；而年代越久远，期间所发生的迁移事件也就越多。在这种情况下，极有可能出现的情况是，其中一人的某个祖先是移民。如此一来，这两个人就不太可能有亲缘关系。迁移率越高，出现这种情况的可能性便会越大。

[44] Langergraber et al., 2011.

[45] Bowles, 2006.

[46] Trivers, 1971.

[47] Axelrod & Hamilton, 1981.

[48] McElreath & Boyd, 2007.

[49] Maskin et al., 2015.

[50] Clutton-Brock, 2009.

[51] Boyd et al., 2014.

[52] 例如，假设当有至少一半的群体成员选择合作时，个体会继续投入战斗，使得某种形式的合作持续下去。在某些情况下，当某个群体中既有合作者又有背叛者时，该群体在进化上反而会相对稳定。不过，只有在群体规模足够小的情况下，才有可能将收益代价比维持在合理范范内。此外，在这一过程中，合作性群体与非合作性群体皆会出现，同时，在某些罕见的情况下，包容性更强的策略会有所增加。当然，这一效应只会出现在规模非常小的群体中。详情请参阅人类学家罗伯托·斯科尔曼·琼（Roberto Schonmann Joan B）和博伊德各自发表于2016年的著作。

[53] Hart & Pilling, 1960.

[54] Marlowe, 2009.

[55] Panchanathan & Boyd, 2004.

[56] E.g., Raihani et al., 2012.

[57] E.g., Panchanathan & Boyd, 2004.

[58] Boyd et al., 2014; Schonmann & Boyd, 2016.

[59] E.g., Marlowe, 2009; Guala, 2012.

[60] Mathew & Boyd 2011, 2014.

[61] Mathew, 2017.

[62] Leimar, 1997.

[63] Richerson & Boyd, 2005; Henrich, 2015.

[64] Reviewed in Richerson et al., 2016.

[65] Holden & Mace, 2003.

[66] Mathew & Perreault, 2015.

[67] Sosis, 2000.

[68] Henrich et al., 2010.

[69] Kelly, 1985.

[70] Soltis et al., 1995.

[71] Henrich & Gil-White, 2001; McElreath et al., 2008.

[72] Boyd & Richerson, 2002.

[73] Stark, 1997.

[74] E.g., Alba & Nee, 2003; Martin, 2005.

[75] E.g., Cronk, 2002.

[76] 群体人口密度依赖性的竞争水平亦非常重要，但为了让论述看起来简单一些，我在此就不做赘述了，请参阅人类学家朱利亚·莱曼（Julia Lehmann）等人发表于2006年的相关评论。

[77] 请参阅生物学家戴维·奎勒（David Queller）撰写的关于这场争论的文章，他清晰且诙谐地解释了为何这两种方法具有等同效用。

[78] E.g., West et al., 2011.

[79] 在基因进化方面，类似过程可能并不重要，因为自然选择的力量要弱得多，且无法维持竞争群体在基因频率上的差异。请参阅博伊德等人发表于2011年的著作。

[80] 群居性昆虫族群的例子未必是反例，正如我们在许多情况下所看到的那样，效率最高的群体通常只包含两个个体，一雌一雄，以繁衍后代为目的。

[81] Boyd et al., 2011b.

[82] West et al., 2011.

[83] Boyd et al., 2011b.

[84] E.g., Boyd & Richerson, 1985, 1990, 2002.

[85] E.g., Gavrilets, 1996.

[86] Zefferman & Mathew, 2015.

[87] Appiah, 2010.

[88] 同上。

[89] E.g., Young, 2001.

[90] E.g., Barton & Rouhani, 1993.

[91] Kaplan et al., 2000.

[92] Clutton-Brock, 2009; Schino & Aureli, 2009.

[93] Sugiyama, 2003.

[94] Mathew et al., 2013.

[95] Meggitt, 1965.

[96] Mathew et al., 2013.

[97] 这是个有趣的进化论解释，说明了人们为何会将这些规范内化，即便这样做会令选择受限。请参阅约瑟夫·亨里奇发表于2015年的著作。

[98] Axelrod & Hamilton, 1981; Lehmann & Keller, 2006.

[99] Nowak & Sigmund, 1993.

[100] Fudenberg et al., 2012.

[101] Panchanathan & Boyd, 2004.

[102] Boyd & Mathew, 2015.

[103] Mathew & Boyd, forthcoming.

[104] E.g., Stevens & Stephens, 2003.

[105] DalBó & Fréchette, 2014.

[106] Gurven, 2004.

[107] Lewis, 2014.

[108] Boyd et al., 2011b.

04 适应不需要洞察力？

[1] 博伊德没有过多论述创新的起源，所以关于这方面的讨论，就暂且先搁置一下。

[2] 事实上，模型参数反映的是外界信息的可靠性和各种行为变量的实际种群发生率，而做出选择的前提是，行动者能够作为这种可靠性和发生率的代表。

[3] 技术不是唯一会对成功率造成影响的因素，所以得出的数据会含有杂乱的信息。

[4] 就近似机制而言，这一点是正确的，但一阶搭便车者所付出的代价和获得的收益，可能会不同于高阶搭便车者的；另外，对一阶搭便车行为的控制也可能会不同于对高阶搭便车行为的。两种控制之间存在系统性差异。

[5] 或许在人们的脑海中已经形成了有关进化的观念，将博伊德和里克森视为一个“超级有机体”，并非不合理。

[6] 非常感谢讲座的组织者邀请我参加讲座并参与讨论；非常感谢博伊德，以及与我一同参与评论的同仁们，这真是一次既愉悦又刺激的经历。

06 适应性、合作性、操控性和敌对性

[1] Boyd & Richerson, 1985; Mathew & Boyd, 2011; and Henrich, 2016.

[2] 生物学家已知的群体选择更青睐个体的非适应性行为，因为它能为处于竞争状态下的群体做出贡献；博伊德所说的文化群体选择则更青睐个体的适应性行为，因为它能维护群体间竞争所选择的不同规范。

07 文化、观念和决策

[1] Scott-Philips, 2014.

[2] Carey, 2011; Pica et al., 2004.

[3] Kimbel & Villmoare, 2016.

[4] Vaesen, 2012.

[5] 围绕该术语，存在不少令人困惑的问题。我与里克森（Boyd & Richerson 1985）将“从众传播”（conformist transmission）定义为一种社会学习规则。我们认为这种规则能够促增适应性特质在人群中出现的频次和在各种环境中的可获得性。有学者对这一推论进行过探

讨，但对于该规则是否真的具有上述效用，目前尚无定论（Henrich, Boyd, 1998; Wakano, Aoki,2007; Perreaykt et al., 2012）。在20世纪50年代所进行的一系列实验中，心理学家使用“从众”这一术语来指代对大众化行为进行模仿的心理倾向，心理学家尼古拉斯·克拉迪埃（Nicolas Claidière）与生物学家安德鲁·怀滕在2012年进行相关研究时也采用了这一术语。根据“从众”一词的定义，我们认为从众传播有助于促增群体中非大众化行为特质的出现。对此，科学哲学家金·斯特林的表述不甚明确，或许是因为我当年尚未提及“从众”一词，不过我猜想，他指的可能是后面这种定义。图1-6所示的ESS传播规则仅适用于出现频次很高且备受欢迎的特质。

[6] Cladière & Sperber, 2007.

[7] 请参阅人类学家罗伯特·博伊德和彼得·里克森发表于1985年和2005年，以及约瑟夫·亨里奇发表于2015年的相关论著。人类学家马克西姆·德雷和博伊德即将发表的一项研究结果表明，在实验室文化模型中，对成功的偏好会导致“高回报”特质的广泛传播，且无须人们了解其间的因果关系。

[8] Boyd & Richerson, 1985; Richerson & Boyd, 2005.

[9] Barrett & Broesch, 2012.

[10] Wertz & Wynn, 2014.

[11] Henrich & McElreath, 2003.

[12] 请参阅博伊德和里克森于1985年提出的与此类偏差有关的进化模型。

[13] Dehaene, 2009.

[14] Knauft, 1985.

[15] Kaplan et al., 1995.

[16] 请参阅里克森和博伊德发表于2005年的著作中的第6章和第7章，更多例证请参阅亨里奇发表于2015年，以及人类学家罗伯特·埃杰顿发表于1992年的著作。

[17] Henrich, 2015.

[18] 因为不了解必要的加工技术，生活在城市里的委内瑞拉人也曾经历过树薯中毒事件。当地人称树薯为“yucca”。请参阅记者尼古拉斯·凯西（Nicholas Casey）发表于2016年12月25日版《纽约时报》的文章《没有食物、药物和片刻的休息：一个饥饿的委内瑞拉男孩之死》（*No Food, No Medicine, No Respite: A Starving Boy's Deathin Venezuela*）。

[19] 文化群体选择可能创造了一种社会环境，在这种环境中，自然选择将作用于偏好狭隘利他主义的遗传变异，且作用的维度与文化变异相同。请参阅里克森和博伊德发表于2005年的著作中的第7章。

[20] Richerson et al., 2016.

[21] 请参阅约瑟夫·亨里奇发表于2015年的著作，在第9章中，他论述了觅食部落的社会规范，并展示了卡拉哈里部落的社会规范的更多细节。

[22] Jaeggi & Gurven, 2013.

[23] Bisin & Verdier, 2001.

[24] Kline et al., 2013.

[25] 请参阅亨里奇发表于2009年的著作。重要的是，需要看到，可靠性之所以无法稳定地持续增强，是因为在生物系统和经济领域里，高代价信息不仅永远无法给信息发送者带来回报，还很容易被他人习得，且无须让信息传递者受益。

[27] Oliver, 1962.

[28] Edgerton, 1985.

[29] Mikhail, 2007.

[30] Zipes, 2012.

01、02

Alba, R., and Nee, V. 2003. *Remaking the American Mainstream: Assimilation and the New Immigration.* Harvard University Press, Cambridge, MA.

Aplin, K. P., Chesser, T., and Have, J. T. 2003. Evolutionary biology of the genus *Rattus*: Profie of an archetypal rodent pest. In *Rats, Mice and People: Rodent Biology and Management,* G. R. Singleton, L. A. Hinds, C. J. Krebs, and D. M. Spratt, eds., 487-498. Australian Centre for International Agricultural Research, Canberra.

Appiah, K. A. 2010. *The Honor Code: How Moral Revolutions Happen.* W. W. Norton, New York.

Arrow, K. 1962. The economic implications of learning by doing. *Review of Economic Studies,*29, 155-173.

Axelrod, R., and Hamilton, W. D. 1981. The evolution of cooperation. *Science*, 211, 1390-1396.

Bamforth, D. P. 1994. Indigenous people, indigenous violence: Precontact warfare on the North American Great Plains. *Man*, 29, 95-115.

Barrett, C., Cosmides, L., and Tooby, J. 2007. The hominid entry into the cognitive niche. In *Evolution of Mind: Fundamental Questions and Controversies*, S. Gangestad and J. Simpson, eds., 241-248. Guilford Press, New York.

Barton, N. H., and Rouhani, S. 1993. Adaptation and the shifting balance. *Genetical Research,* 61, 57-74.

Basalla, G. 1988. *The Evolution of Technology.* Cambridge University Press, Cambridge, UK.

Beppu, A., and Griffiths, T. L. 2009. Iterated learning and the cultural ratchet. *Proceedings of the 31st Annual Conference of the Cognitive Science Society.*

Binford, L. 2001. *Constructing Frames of Reference.* University of California Press, Berkeley.

Birch, S., and Okasha, S. 2014. Kin selection and its critics. *Bio Science,*

Bisin, A., and Verdier, T. 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory,* 97, 298-319.

Bliege-Bird, R., Bird, D. W., Codding, B. F., Parker, C. H., and Jones, J. H. 2009. The “fire stick farming” hypothesis: Australian Aboriginal foraging strategies, biodiversity, and anthropogenic fire mosaics. *Proceedings of the National Academy of Sciences (USA)*, 105, 14796-14801.

Boomsma, J. J., Beekman, M., Cornwallis, C. K., Griffin, A. S., Holman, L., Hughes, W. O. H., Keller, L., Oldroyd, B. P., and Ratnieks, F. L. W. 2011. Only full sibling families evolved eusociality. *Nature*,

Bowles, S. 2006. Group competition, reproductive leveling, and the evolution of human altruism. *Science*, 314, 1569-1572.

Boyd, R., and Mathew, S. 2015. Thirdparty monitoring and sanctions aid the evolution of language. *Evolution and Human Behavior*, 36, 475-479.

Boyd, R., and Richerson, P. J. 1985. *Culture and the Evolutionary Process*. University of Chicago Press, Chicago.

——. 1987. The evolution of social learning: The effects of spatial and temporal variation. In *Social Learning: Psychological and Biological Perspectives*, T. R. Zentall and B. G. Galef, eds. Psychology Press, New York.

——. 1990. Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, 145, 331-342.

——. 1995. Why does culture increase human adaptability? *Ethology and Sociobiology*, 16,125-143.

——. 1996. Why culture is common but cultural evolution is rare. *Proceedings of the British Academy*, 88, 73-93.

——. 2002. Group beneficial norms spread rapidly in a structured population. *Journal of Theoretical Biology*, 215, 287-296.

Boyd, R., Richerson, P. J., and Henrich, J. 2011a. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences (USA)*, 108, 10918-10925.

——. 2011b. Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioral Ecology and Sociobiology*, 65, 431-444.

Boyd, R., Schonmann, R. H., and Vicente, R. 2014. Hunter-gatherer population structure and the evolution of contingent cooperation. *Evolu-tion and Human Behavior,* 35, 219-227.

Boyer, P. 2001. *Religion Explained: The Evolutionary Origins of Religious Thought*. Basic Books, New York.

Brink, J. W. 2005. Inukshuk: Caribou drive lanes on southern Victoria Island, Nunavut, Canada. *Arctic Anthropology*, 42, 1-28.

——. 2008. *Imagining Head-Smashed-In: Aboriginal Buffalo Hunting on the Northern Great Plains*. Athabasca University Press, Edmonton, AB, Canada.

——. 2013. The Barnett site: A stone drive lane communal pronghorn trap on the Alberta Plains, Canada. *Quaternary International*, 297, 24-35.

Broesch, J. 2010. *Cultural Transmission and the Role of Evolutionary Forces: Empirical Evidence for Adaptive Biases in Learning and Cultural Transmission.* PhD diss., Emory University.

Cathcart, M. 2014. *Starvation in a Land of Plenty: Wills' Diary of the Fateful Burke and Wills Expedition.* National Library of Australia, Kindle ed.

Clarke, P. A. 2012. *Australian Plants as Aboriginal Tools.* Rosenberg Publishing, Kindle ed.

Clastres, P. 1972. The Guayaki. In *Hunters and Gatherers Today,* M. B. Bicchieri, ed. Holt, Rinehart and Winston, New York.

Clutton-Brock, T. 2009. Cooperation between non-kin in animal societies. *Nature,* 462, 51-57.

Collias, E., and Collias, N. 1964. The development of nest-building behavior in a weaverbird. *The Auk,* 81, 42-52.

——. 1973. Further studies on development of nest-building behavior in a weaverbird(*Ploceus cucullatus*). *Animal Behavior,* 21, 371-382.

Cosmides, L., and Tooby, J. 2006. Evolutionary psychology: A primer. Center for Evolutionary

Psychology, University of California, Santa Barbara.

Couzin, I. D., and Franks, N. R. 2003. Self-organized lane formation and optimized traffic flow in army ants. *Proceedings of the Royal Society B,* 270, 139-146.

Cronk, L. 2002. From true Dorobo to Mukogodo Maasai: Contested ethnicity in Kenya. *Ethnology,* 41, 27-49.

Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T., and Mace, R. 2010. Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature,* 467, 801-804.

Dal Bó, P., and Fréchette, G. R. 2014. On the determinants of cooperation in infinitely repeated games: A survey. SSRN 2535963.

Darwent, J., and Darwent, C. M. 2014. Scales of violence across the North American Arctic. In *Violence and Warfare among Hunter-Gatherers,* M. W. Allen and T. L. Jones, eds., 149-167. Left Coast Press, Walnut Creek, CA.

Dawkins, R. 1986. *The Blind Watchmaker.* W. W. Norton, New York.

Dean, L. a., Val, G. L., Laland, K. N., Flynn, E., and Kendal, R. L. 2014. Human cumulative culture: A comparative perspective. *Biological Reviews,* 89, 284-301.

Dehaene, S. 2009. *Reading and the Brain.* Penguin Viking, New York.

Denham, T., Fullagar, R., and Head, L. 2009. Plant exploitation on Sahul: From colonisation to the emergence of regional specialisation during the Holocene. *Quaternary International,*202, 29-40.

Derex, M., and Boyd. R. 2015. The foundations of the human cultural niche. *Nature Communications.*

Earl, J. W., and McCleary, B. V. 1994. Mystery of the poisoned expedition. *Nature,* 368, 683-684.

Enquist, M., Ghilanda, S., Jarrick, A., and Wachtmeister, C-A. 2008. Why does culture increase exponentially? *Theoretical Population Biology,* 74, 46-55.

Fitzpatrick, K. 2008. The Burke and Wills Expedition and the Royal Society of Victoria. *Historical Studies: Australia and New Zealand,* 10(40), 470-478.

Franks, N. R. 1986. Teams in social insects: Group retrieval of prey by army ants (*Eciton burchelli,* Hymenoptera: Formicidae). *Behavioral Ecology and Sociobiology,* 18, 425-429.

Friesen, T. M. 2013. The impact of weapon technology on caribou drive system variability in the prehistoric Canadian Arctic. *Quaternary International,* 297, 13-23.

Fry, D. 2006. *The Human Potential for Peace: An Anthropological Challenge to Assumptions about War and Violence.* Oxford University Press, Oxford, UK.

Fudenberg, D., Rand, D. G., and Dreber, A. 2012. Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review,* 102, 720-749.

Gat, A. 2015. Proving communal warfare among hunter-gatherers: The quasi-Rousseauan error. *Evolutionary Anthropology,* 24, 111-126.

Gavrilets, S. 1996. On phase three of the shifting-balance theory. *Evolution,* 50, 1034-1041.

Gergely, G., and Csibra, G. 2006. Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In *Roots of Human Sociality: Culture, Cognition, and Human Interaction,* N. J. Enfield and S. C. Levenson, eds., 229-255. Berg Publishers,

Oxford, UK.

Gingrich, P. D. 1983. Rates of evolution: Effects of time and temporal scaling. *Science,* 222,159-161.

Gould, J. L., and Gould, C. G. 2012. *Animal Architects: Building and the Evolution of Intelligence.* Basic Books, New York.

Greenhouse, S. 2014. More workers are claiming “wage theft.” *New York Times,* August 31.

Guala, F. 2012. Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioural and Brain Sciences,* 35, 1-59.

Gurven, M. 2004. Reciprocal altruism and food sharing decisions among Hiwi and Ache huntergatherers. *Behavioral Ecology and Sociobiology,* 56, 366-380.

Hailey, K., and Fessler, D. 2005. Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior,* 26, 245-256.

Hämäläinen, P. 2008. *The Comanche Empire.* Yale University Press, New Haven, CT.

Hansell, M. 2005. *Animal Architecture.* Oxford University Press, New York.

Hart, C. W. M., and Pilling, A. R. 1960. *The Tiwi of North Australia.* Holt, Rinehart and Winston, New York.

Henrich, J. 2004. Demography and cultural evolution: Why adaptive cultural processes produced maladaptive losses in Tasmania. *American Antiquity,* 69, 197-218.

——. 2006. Understanding cultural evolutionary models: A reply to Read's critique. *American Antiquity,* 71, 771-782.

——. 2009. The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior,*30, 244-260.

——. 2015. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter.* Princeton University Press, Princeton, NJ.

Henrich, J., and Boyd, R. 2001. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology,* 208, 79-89.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogo, C., Marlowe, F., Tracer, D., and Ziker, J. 2010. Markets, religion, community size, and the evolution of fairness and punishment. *Science,* 327, 1480-1484.

Henrich, J., and Gil-White, F. J. 2001. The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior,* 22, 165-196.

Henrich, J., and Henrich, N. 2010. The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings of the Royal Society B,* 277, 3715-3724.

Hill, K. 2002. Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature,* 13, 105-128.

Hill, K., Barton, M., and Hurtado, A. M. 2009. The origins of human uniqueness: The

evolution of characters underlying behavioral modernity. *Evolutionary Anthropology,* 18, 187-200.

Hill, K. R., Walker, R. S., Božičević, M., Eder, J., Headland, T., Hewlett, B., Hurtado, A. M., Marlowe, F., Wiessner, P., and Wood, B. 2011. Coresidence patterns in hunter-gatherer societies show unique human social structure. *Science,* 331, 286-289.

Hill, K. R., Wood, B. M., Baggio, J., Hurtado, A. M., and Boyd, R. 2014. Hunter-gatherer inter-band interaction rates: Implications for cumulative culture. *PLoS One.*

Holden, C. J., and Mace, R. 2003. Spread of cattle led to the loss of matrilineal descent in Africa: A coevolutionary analysis. *Proceedings of the Royal Society B*.

Howitt, A. W., Burke, R. O., King, J., Mueller, F., Wills, W. J., Wright, W., Brahé, W., Ligon, C. W. V., MacAdam, J., Smith, J., and Archer, Mr. 1862. *Exploring Expedition from Victoria to the Gulf of Carpentaria, under the Command of Mr. Robert O'Hara Burke.*

Iwaniuk, A. N., and Arnold, K. E. 2004. Is cooperative breeding associated with bigger brains? A comparative test in the Corvida (Passeriformes). *Ethology,* 110, 203-220.

Jacklyn, P. M. 1992. “Magnetic” termite mound surfaces are oriented to suit wind and shade conditions. *Oecologica,* 91, 385-395.

James, S. R., and Graziani, S. 1992. California Indian warfare. UC Berkeley Digital Assets.

Johnson, J. A. 1987. Dominance rank in juvenile olive baboons, Papio anubis: The infuence of gender, size, maternal rank and orphaning. *Animal Behaviour,* 35, 1694-1708.

Johnson, S. 2010. *Where Good Ideas Come From: The Natural History of Innovation.* Riverhead Hardcover, New York.

Jordan, P. 2015. *Technology as Social Tradition: Cultural Transmission among Hunter Gatherers.* University of California Press, Berkeley.

Kaplan, H., Hill, K., Lancaster, J., and Hurtado, A. M. 2000. A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology,* 9, 156-185.

Kay, E. H., and Hoekstra, H. E. 2008. Rodents. *Current Biology,* 18, R408-R410.

Kelly, R. 1985. *The Nuer Conquest.* University of Michigan Press, Ann Arbor.

Klein, R. 2009. *The Human Career.* 3rd ed. University of Chicago Press, Chicago.

Kline, M., and Boyd, R. 2010. Population size predicts technological complexity in Oceania.

Proceedings of the Royal Society B, 277, 2559-2564.

Kroeber, A. L. (1925) 1976. *The Handbook of California Indians.* Dover, New York.

Lane, N. 2009. *Life Ascending.* Profile Books, London.

Langergraber, K., Schubert, G., Rowney, C., Wrangham, R., Zommers, Z., and Vigilant, L. 2011. Genetic differentiation and the evolution of cooperation in chimpanzees and humans. *Proceedings of the Royal Society B*.

LeBlanc, S. 2014. Forager warfare and our evolutionary past. In *Violence and Warfare among Hunter-Gatherers,* M. W. Allen and T. L. Jones, eds. 149-167. Left Coast Press, Walnut Creek, CA.

Lehmann, L., and Keller, L. 2006. The evolution of cooperation and altruism: A general framework and a classification of models. *Journal of Evolutionary Biology,* 19, 1365-1376.

Lehmann, L., Perrin, N., and Rousset, F. 2006. Population demography and the evolution of helping behaviors. *Evolution,* 60, 1137-1151.

Leimar, O. 1997. Repeated games: A state space approach. *Journal of Theoretical Biology,*184, 471-498.

Lewis, J. 2014. Egalitarian social organization: The case of the Mbendjele Ba Yaka. In *Hunter-Gatherers of the Congo Basin,* B. S. Hewlett, ed., Kindle location 5800. Transaction Publishers, Kindle ed.

Lukas, D., and Clutton-Brock, T. 2012a. Cooperative breeding and monogamy in mammalian societies. *Proceedings of the Royal Society B,* 279, 2151-2156.

——. 2012b. Life histories and the evolution of cooperative breeding in mammals. *Proceedings of the Royal Society B,* 279, 4065-4070.

Manson, J., and Wrangham, R. 1991. Intergroup aggression in chimpanzees and humans. *Current Anthropology,* 32, 369-377.

Marlowe, F. 2009. Hadza cooperation: Second party punishment, yes; third party punishment, no. *Human Nature,* 20, 417-430.

Martin, P. 2005. Migrants in the global labor market. Global Commission on International Migration Working Paper.

Maskin, E., Harrison, L., and Yasin, Y. 2015. Culture, cooperation, and repeated games. In *Culture Matters in Russia and Everywhere,* L. Harrison and E. Yasin, eds. Lexington Press, Lanham, MD.

Mathew, S. 2017. How the second-order free rider problem is solved in a smallscale society. *American Economic Review,* 107, 1-4.

Mathew, S., and Boyd, R. 2011. Punishment sustains largescale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences (USA),* 108, 11375-11380.

——. 2014. The cost of cowardice. *Evolution and Human Behavior,* 35, 58-64.

——. Forthcoming. Adjudication and perception errors.

Mathew, S., Boyd, R., and van Veelen, M. 2013. Human cooperation among kin and close associates may require enforcement of norms by third parties. In *Cultural Evolution,* P. J. Richerson and M. Christiansen, eds. Strüngmann Forum Report 12, J. Lupp, series ed. MIT Press, Cambridge, MA.

Mathew, S., and Perreault, C. 2015. Behavioural variation in 172 small-scale societies indicates that social learning is the main mode of human adaptation. *Proceedings of the Royal Society B.*

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., and Waring, T. 2008. Beyond existence and aiming outside the laboratory: Estimating frequency-dependent and payoff-biased social learning strategies. *Philosophical Transactions of the Royal Society B,* 363, 3515-3528.

McElreath, R., and Boyd, R. 2007. *Mathematical Models of Social Evolution: A Guide for the Perplexed.* University of Chicago Press, Chicago.

Meggitt, M. J. 1965. *Desert People: A Study of the Walbiri Aborigines of Central Australia.* University of Chicago Press, Chicago.

Moorehead, A. 1963. *Cooper's Creek.* Dell, New York.

Morgan, J. (1852) 2002. *The Life and Adventures of William Buckley.* Text Publishing, Melbourne, Australia.

Morgan, T. H., Uomini, N., Rendell, L. E., Chouinard-Thuly, L., Street, S., and Lewis, H. 2015. Experimental evidence for the coevolution of hominin tool-making teaching and language. *Nature Communications* (January), 6.

Nowak, M., Page, K., and Sigmund, K. 2000. Fairness versus reason in the ultimatum game. *Science,* 289, 1773-1775.

Nowak, M., and Sigmund, K. 1993. A strategy of winstay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature,* 364, 56-58.

Nowak, R. M. 1999. *Walker's Mammals of the World.* Johns Hopkins University Press, Baltimore.

Neiman, F. D. 1995. Stylistic variation in evolutionary perspective: Infer-ences from decorative diversity and interassemblage distance in Illinois Woodland ceramic assemblages. *American Antiquity,* 60, 7-36.

O'Shea, J., Lemke, A. K., and Reynolds, R. G. 2013. “Nobody knows the way of the caribou”: *Rangifer* hunting at 45° north latitude. *Quaternary International,* 297, 36-44.

Panchanathan, K., and Boyd, R. 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature,* 432, 499-502.

Perreault, C. 2012. The pace of cultural evolution. *PLoS ONE.*

Perreault, C., Moya, C., and Boyd, R. 2012. A Bayesian approach to the evolution of social learning. *Evolution and Human Behavior,* 33, 449-459.

Petroski, H. 1992a. *The Evolution of Useful Things: How Everyday Artifacts From Forks and Pins to Paper Clips and Zippers Came to Be as They Are.* Vintage Books, New York.

——. 1992b. *The Pencil.* Knopf, New York.

Pinker, S. 2010. The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences (USA),* 107, 8993-8999.

Powell, A., Shennan, S., and Thomas, M. G. 2009. Late Pleistocene demography and the appearance of modern human behavior. *Science,* 324, 1298-1301.

Raihani, N., Thornton, A., and Bshary, R. 2012. Punishment and cooperation in nature. *Trends in Ecology and Evolution,* 27, 288-295.

Richerson, P. J., Bell, A., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E., Narr, N., Newson, L., Ross, C., Smaldino, P., Waring, T., and Zefferman, M. 2016. Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioural and Brain Sciences,* 39, 1-68.

Richerson, P. J., and Boyd, R. 2005. *Not by Genes Alone.* University of Chicago Press, Chicago.

Romer, P. 1993. Endogenous technological change. *Journal of Political Economy,* 98, 71-102.

Scherjon, F., Bakels, C., MacDonald, K., and Roebroeks, W. 2015. Burning the land: An ethnographic study of off-site fire use by current and historically documented foragers and implications for the interpretation of past fire practices in the landscape. *Current Anthropology,* 56, 299-326.

Schino, G., and Aureli, F. 2009. Reciprocal altruism in primates: Partner choice, cognition, and emotions. *Advances in the Study of Behavior,* 39, 45-69.

Schonmann, R. H., and Boyd, R. 2016. A simple rule for the evolution of contingent cooperation in large groups. *Philosophical Transactions of the Royal Society B,* dx.

Seabright, P. 2010. *In the Company of Strangers.* Princeton University Press, Princeton, NJ.

Sharif, A., Norenzayan, A., and Henrich, J. 2010. The birth of high gods: How the cultural evolution of supernatural policing influenced the evolution of complex cooperative human societies paving the way for civilization. In *Evolution, Culture and the Human Mind,* M. Schaller, A. Norenzayan, S. J. Heine, T. Yamagishi, and T. Kameda, eds., 119-136. Taylor and Francis, New York.

Shennan, S. 2001. Demography and cultural innovation: A model and its implications for the emergence of modern human culture. *Cambridge Archaeological Journal,* 11, 5-16.

Smith, B. D. 2011. General patterns of niche construction and the management of "wild" plant and animal resources by smallscale preindustrial societies. *Philosophical Transactions of the Royal Society B,* 366, 836-848.

Smith, M. W. 1938. The war complex of the Plains Indians. *Proceedings of the American Philosophical Society,* 78, 425-464.

Soltis, J., Boyd, R., and Richerson, P. J. 1995. Can group functional behaviors evolve by cultural group selection? An empirical test. *Current Anthropology,* 36, 473-494.

Sosis, R. 2000. Religion and intragroup cooperation: Preliminary results of a comparative analysis of utopian communities. *Cross Cultural Research,* 34, 70-88.

Stark, R. 1997. *The Rise of Christianity: How the Obscure, Marginal Jesus Movement Became the Dominant Religious Force in the Western World in a Few Centuries.* Harper Collins, San Francisco.

Steinmetz, P. H. R., Kraus, J. E. M., Larroux, C., Jörg, U., Hammel, J. U., Amon-Hassenzahl, A., Houliston, E., Wörheide, G., Nickel, M., Degnan, B. M., and Technau, U. 2012. Independent evolution of striated muscles in cnidarians and bilaterians. *Nature,* 487,231-236.

Sterelny, K. 2012. *The Evolved Apprentice: How Evolution Made Humans Unique.* MIT Press, Cambridge, MA.

Stevens, J. R., and Stephens, D. W. 2003. The economic basis of cooperation: Tradeoffs between selfishness and generosity. *Behavioral Ecology,* 15, 255-261.

Steward, J. 1933. Ethnography of the Owens Valley Paiute. *University of California Publications in American Archaeology Ethnology,* 33, 233-350.

Sugiyama, L. S. 2003. Illness, injury, and disability among Shiwiar forager-horticulturalists: Implications of health-risk buffering for the evolution of human life history. *American Journal of Physical Anthropology,* 123, 371-389.

Sutton, M. Q. 2014. Warfare and expansion: An ethnohistoric perspective on the Numic spread. In *Violence and Warfare among Hunter-Gatherers,* M. W. Allen and T. L. Jones, eds.,149-167. Left Coast Press, Walnut Creek, CA.

Swezey, S. L., and Heizer, R. H. 1977. Ritual management of salmonid fish resources in California. *Journal of California Anthropology,* 4, 429.

Tebbich, S., Taborsky, M., Fessl, B., and Blomqvist, D. 2001. Do woodpecker finches acquire tool-use by social learning? *Proceedings of the Royal Society B,* 268, 2189-2193.

Tennie, C., Call, J., and Tomasello, M. 2009. Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B,* 364, 2405-2415.

Tooby, J., and DeVore, I. 1987. The cognitive niche. In *Primate Models of Hominid Behavior,* W. Kinzey, ed., 183-237. SUNY Press, New York.

Trivers, R. L. 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology,* 46,35-57.

Wadley, L., Hodgskiss, T., and Grant, M. 2009. Implications for complex cognition from the hafting of tools with compound adhesives in the Middle Stone Age, South Africa. *Proceedings of the National Academy of Sciences (USA),* 106, 9590-9594.

West, S. A., El Moden, C., and Gardner, A. 2011. Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior,* 32, 231-262.

Whiten, A., Horne, V., and de Waal, F. B. M. 2005. Conformity to cultural norms of tool use in chimpanzees. *Nature,* 437, 737-740.

Wilke, P. 2013. The Whisky Flat pronghorn trap complex, Mineral County, Nevada, western United States: Preliminary report. *Quater-nary International,* 297, 79-92.

Wood, B., and Gilby, I. Forthcoming. Hunting and meat sharing by chimpanzees, humans, and our common ancestor. In *Chimpanzees and Human Evolution,* M. Muller, R. Wrangham, and D. Pilbeam, eds. Harvard University Press, Cambridge, MA.

Young, P. 2001. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton University Press, Princeton, NJ.

Zefferman, M. R., and Mathew, S. 2015. An evolutionary theory of large-scale human warfare: Group-structured cultural selection. *Evolution-ary Anthropology,* 24, 50-61.

03

Boyd, R., and P. J. Richerson. 1993. Rationality, imitation, and tradition. In *Nonlinear Dynamics and Evolutionary Economics.* R. Day and P. Chen, eds., 131-149. Oxford University Press, New York.

Burke, E. (1790) 1910. *Reflections on the Revolution in France.* J. M. Dent and Sons, London.

Cavalli-Sforza, L. L., and M. W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton University Press, Princeton, NJ.

Hayek, F. A., 1952. *The Counter-Revolution of Science.* Free Press, Glencoe, IL.

——. 1984a. Dr. Bernard Mandeville. In *The Essence of Hayek,* C. Nishiyama and K. R. Leube, eds., 176-194. Hoover Institution Press, Stanford, CA.

——. 1984b. *The Essence of Hayek.* C. Nishiyama and K. R. Leube, eds. Hoover Institution Press, Stanford, CA.

Lescak, E. A., Bassham, S. L., Catchen, J., Gelmond, O., Sherbick, M. L., von Hippel, F. A., and Cresko, W. A. 2015. Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proceedings of the National Academy of Sciences (USA),* 112, E7204-E7212.

Orr, H. A. 2009. Darwin and Darwinism: The (alleged) social implications of *The Origin of*

Species. Genetics, 183, 767-772.

——. 2015. The biology of being good to others (review of David Sloan Wilson's *Does Altruism Exist?) New York Review of Books* (March), 1927-1929.

Perreault, C., Moya, C., and Boyd, R. 2012. A Bayesian approach to the evolution of social learning. *Evolution and Human Behavior,* 33, 449-459.

Price, G. R. 1970. Selection and covariance. *Nature,* 227(5257), 520-521.

——. 1972. Extension of covariance selection mathematics. *Annals of Human Genetics,* 35,485-490.

Secord, J. A., ed. 2008. *Charles Darwin: Evolutionary Writings.* Oxford University Press, Oxford, UK.

Subiaul, F. 2007. The imitation faculty in monkeys: Evaluating its fea-tures, distribution and evolution. *Journal of Anthropological Sciences,* 85, 35-62.

Tooby, J., and DeVore, I. 1987. The cognitive niche. In *Primate Models of Hominid Behavior,* W. Kinsey, ed., 183-237. SUNY Press, New York.

04

Alvard, M., and Nolin, D. 2002. Rousseau's whale hunt? Coordination among big game hunters. *Current Anthropology,* 43(4), 533-559.

Edgerton, R. B. 1992. *Sick Societies: Challenging the Myth of Primitive Harmony.* Free Press, New York.

Ewald, P. W. 1994. *Evolution of Infectious Disease.* Oxford University Press, Oxford, UK.

Harris, M. 1985. *Good to Eat: Riddles of Food and Culture.* Simon and Schuster, New York.

Henrich, J. 2004. Demography and cultural evolution: Why adaptive cultural processes produced maladaptive losses in Tasmania. *American Antiquity,* 69(2), 197-221.

——. 2016. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species and Making Us Smarter.* Princ-eton University Press, Princeton, NJ.

Henrich, J., and Henrich, N. 2010. The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings of the Royal Society B,* 277, 3715-3724.

Hewlett, B., Fouts, H., Boyette, A., and Hewlett, B. 2011. Social learning among Congo Basin hunter-gatherers. *Philosophical Transactions of the Royal Society of London B,* 366(1567), 1168-1178.

Heyes, C. 2012. Grist and mills: On the cultural origins of cultural learning. *Philosophical Transactions of the Royal Society of London B,* 367(1599), 2181-2191.

Hiscock, P. 2014. Learning in lithic landscapes: A reconsideration of the hominid “toolmaking” niche. *Biological Theory,* 9(1), 27-41.

Liebenberg, L. 1990. *The Art of Tracking and the Origin of Science.* David Philip, Claremount, South Africa.

——. 2008. The relevance of persistence hunting to human evolution. *Journal of Human Evolution,* 55(6), 1156-1159.

——. 2013. *The Origin of Science: On the Evolutionary Roots of Science and Its Implications for Self-Education and Citizen Science.* Cyber Tracker, Capetown, South Africa.

McBrearty, S., and Brooks, A. 2000. The revolution that wasn't: A new interpretation of the origin of modern human behavior. *Journal of Human Evolution,* 39(5), 453-563.

Morin, O. 2016. *How Traditions Live and Die.* Oxford University Press, Oxford, UK.

Nichols, S. 2004. *Sentimental Rules: On the Natural Foundations of Moral Judgment.* Oxford University Press, New York.

Odling-Smee, J., Laland, K., and Feldman, M. 2003. *Niche Construction: The Neglected Process in Evolution.* Princeton University Press, Princeton, NJ.

Powell, A., Shennan, S., and Thomas, M. 2009. Late Pleistocene demog-raphy and the appearance of modern human behavior. *Science,* 324 (June 5), 298-1301.

Richerson, P., and Boyd, R. 2013. Rethinking paleoanthropology: A world queerer than we had supposed. In *The Evolution of Mind, Brain and Culture,* G. Hatfield and H. Pittman, eds., 263-302. University of Pennsylvania Press, Philadelphia.

Shaw-Williams, K. 2014. The social trackways theory of the evolution of human cognition. *Biological Theory,* 9(1), 16-26.

Sterelny, K. 2012. *The Evolved Apprentice.* MIT Press, Cambridge, MA.

Stout, D. 2002. Skill and cognition in stone tool production: An ethnographic case study from Irian Jaya. *Current Anthropology,* 43(5), 693-722.

05

Bauer, M., Cassar, A., Chytilova, J., and Henrich, J. 2014. War's enduring effects on the development of egalitarian motivations and in-group biases. *Psychological Science,* 25(1),47-57.

Boyd, R., and Richerson, P. J. 1985. *Culture and the Evolutionary Process.* University of Chicago Press, Chicago.

Dawkins, R. 1976. *The Selfish Gene.* Oxford University Press, Oxford, UK.

Gneezy, A., and Fessler, D. M. T. 2012. Conflict, sticks and carrots: War increases prosocial punishments and rewards. *Proceedings of the Royal Society B,* 279(1727), 219-223.

Mace, R. 2014. Human behavioral ecology and its evil twin. *Behavioral Ecology,* 25(3), 443-449.

Silva, A. S., and Mace, R. 2014. Cooperation and conflict: Field experiments in Northern Ireland. *Proceedings of the Royal Society B,* 281(1792).

——. 2015. Inter-group conflict and cooperation: Field experiments before, during and after sectarian riots in Northern Ireland. *Frontiers in Psychology,* 6.

Ting, J., Zheng, X.-D., He, Q.-Q., Wu, J.-J., Mace, R., and Tao, Y. 2016. Kinship as a frequency-dependent strategy. *Royal Society Open Science* 3(2).

06

Booker, C. 2006. *The Seven Basic Plots: Why We Tell Stories.* Bloomsbury, London.

Boyd, R., and Richerson, P. J. 1985. *Culture and the Evolutionary Process.* University of Chicago Press, Chicago.

Campbell, J. 1949. *The Hero with a Thousand Faces.* Princeton University Press, Princeton, NJ.

Graça da Silva, S., and Tehrani, J. J. 2016. Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales. *Royal Society Open Science,* 3, 150645.

Henrich, J. 2016. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species and Making Us Smarter.* Princeton University Press, Princeton, NJ.

Mathew, S., and Boyd, R. 2011. Punishment sustains largescale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences (USA),* 108, 11375-11380.

07

Barrett, H. C., and Broesch, J. 2012. Prepared social learning about dangerous animals in children. *Evolution and Human Behavior,* 33, 499-508.

Bisin, A., and Verdier, T. 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory,* 97, 298-319.

Boyd, R., and Richerson, P. J. 1985. *Culture and the Evolutionary Process.* University of Chicago Press, Chicago.

Carey, S. 2011. *The Origin of Concepts.* Oxford University Press, Oxford, UK. Kindle ed.

Claidière, N., and Sperber, D. 2007. The role of attraction in cultural evolution. *Journal of Cognition and Culture,* 7, 89-111.

Claidière, N., and Whiten, A. 2012. Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin,* 138, 126-145.

Dehaene, S. 2009. *Reading in the Brain.* Viking, New York.

Derex, M., and Boyd, R. Forthcoming. What do people learn when they study socially?

Edgerton, R. 1985. *Rules, Exceptions and Social Order.* University of California Press, Berkeley.

——. 1992. *Sick Societies: Challenging the Myth of Primitive Harmony.* Free Press, New York.

Feil, D. K. 1987. *The Evolution of Highland Papua New Guinea Societies.* Cambridge University Press, Cambridge.

Henrich, J. 2009. The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior,* 30, 244-260.

——. 2015. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter.* Princeton University Press, Princeton, NJ.

Henrich, J., and Boyd, R. 1998. The evolution of conformist transmis-sion and the emergence of between-group differences. *Evolution and Human Behavior,* 19, 215-242.

Henrich, J., and McElreath, R. 2003. The evolution of cultural evolution. *Evolutionary Anthropology,* 12, 123-135.

Hill, K., and Hurtado, A. M. 2009. Cooperative breeding in South American hunter-gatherers. *Proceedings of the Royal Society B,* 276, 3863-3870.

Jaeggi, A., and Gurven, M. 2013. Natural cooperators: Food sharing in humans and other primates. *Evolutionary Anthropology,* 22, 186-195.

Kaplan, H. S., Lancaster, J. B., Johnson, S. E., and Bock, J. 1995. Does observed fertility maximize fitness among New Mexican men? *Human Nature,* 6, 325-360.

Kimbel, W. H., and Villmoare, B. 2016. From *Australopithecus to Homo:* The transition that wasn't. *Philosophical Transactions of the Royal Society B,* 371, 20150248.

Kline, M. A., Boyd, R., and Henrich, J. 2013. Teaching and the life history of cultural transmission in Fijian villages. *Human Nature,* 24, 351-374.

Knauft, B. 1985. *Good Company and Violence: Sorcery and Social Action in a Lowland New Guinea Society.* University of California Press, Berkeley.

Mikhail, J. 2007. Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Science,* 11, 143-152.

Oliver, S. C. 1962. *Ecology and Cultural Continuity as Contributing Factors in the Social Organization of the Plains Indians.* University of California Publications in American Archaeology and Ethnology 48, Berkeley.

Perreault, C., Moya, C., and Boyd, R. 2012. A Bayesian approach to the evolution of social learning. *Evolution and Human Behavior,* 33, 449-459.

Pica, P., Lemer, C., Izard, V., and Dehaene, S. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science,* 306, 499-503.

Richerson, P. J., Baldini, R., Bell, A. V., Demps, K., Frosta, K., Hillis, V., Mathew, S., Newton, E. K., Naara, E., Newson, L., Ross, C., Smaldino, P. E., Waringa, T. M., and Zefferman, M. 2016. Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioural and Brain Sciences,* 39, 1-68.

Richerson, P. J., and Boyd, R. 2005. *Not by Genes Alone.* University of Chicago Press, Chicago.

Scott-Philips, T. 2014. *Speaking Our Minds: Why Human Communication Is Different, and How Language Evolved to Make It Special.* Palgrave MacMillan, Basingstoke, UK. Kindle ed.

Seabright, P. A. 2010. *In the Company of Strangers: A Natural History of Economic Life.* Rev. ed. Princeton University Press, Princeton, NJ.

Vaesen, K. 2012. The cognitive bases of human tool use. *Behavioural and Brain Sciences,* 35,203-218.

Wakano, J., and Aoki, K. 2007. Do social learning and conformist bias coevolve? Henrich and Boyd revisited. *Theoretical Population Biology,* 72, 504-512.

Wertz, A. E., and Wynn, K. 2014. Selective social learning of plant edibility in 6- and 18-month-old infants. *Psychological Science,* 25, 874-882.

Wiessner, P. W., Tumu, A., and Pupu, N. 1998. *Historical Vines: Enga Networks of Exchange, Ritual, and Warfare in Papua New Guinea.* Smithsonian Institution Press, Washington, DC.

Zipes, J. 2013. *The Irresistible Fairy Tale: The Social and Cultural History of a Genre.* Princeton University Press, Princeton, NJ.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读App玩转指南

湛庐阅读App结构图：



三步玩转湛庐阅读App：



App获取方式：

安卓用户前往各大应用市场、苹果用户前往App Store 直接下载“湛庐阅读”App，与最聪明的人共同进化！

使用App扫一扫功能，遇见书里书外更大的世界！



延伸阅读

《人类起源的故事》

◎ 传统的考古学、语言学构建的人类演化史留下了大片大片的空白。幸好，这已经成为过去时。古人类DNA测序正在掀起一场革命，重新谱写一曲50万年的人类迁徙演化之歌。

◎ 大卫·赖克从古人类遗骨中提取DNA，从基因层面还原了人类祖先的面貌。原来，我们的先祖在地球上已经上演了几百万年的“权力的游戏”：在任何一片大陆上，人群都经历了多次毁灭与更迭；所有当代人的祖先都拥有一段复杂难辨的混血史，没有人是“纯种”；千百万年来的种族、性别、阶层不平等在我们每个人的DNA里都留下了深深的刻痕。

◎ 面对DNA序列分析这种最先进可靠的技术，我们以前所以为的所有关于人类演变的知识，凡是跟这本书所代表的当前科学理解不一样的，都只能改写。因此，这场革命给了两个问题迄今最为清晰的答案：我们是谁，我们从哪里来。





《当下的启蒙》

◎ 当代最伟大思想家史蒂芬·平克全面超越自我的巅峰之作，一部关于人类进步的英雄史诗。

◎ 通过75幅震撼的图表，平克论证人类的寿命、健康、食物、和平、知识、幸福等都呈向上趋势，这种趋势不仅限于西方，而是遍及全世界。这是启蒙运动的礼物——理性、科学和人文主义促进了人类的进步。

◎　比尔·盖茨最喜爱的一本书。理查德·道金斯心中的诺贝尔文学奖作品。尤瓦尔·赫拉利2018年最爱的书之一。



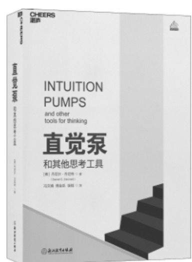


《直觉泵和其他思考工具》

◎　享誉世界的哲学泰斗丹尼尔·丹尼特，融通计算机科学、心理学、神经科学、语言学、人工智能，倾囊相授他一生至今所搜集的各种好用的思考工具。这本书诞生于大学新生的课堂，力图做到“人人能懂”。

◎　使用大量方便的、辅助性的思考工具，去拓展想象力、保持专注力，让我们妥当、优雅地思考真正的难题。利用各种思考工具，让你拨开各种思想的层层迷雾，你会发现，那么多明摆着的观点其实根本就不是那么“明摆着”的。

◎　“直觉泵”是很有用的思考工具，作为一种思想实验，一个好的直觉泵比任何一种论证和分析都更为有力。这本书不仅带你去检验不合格的直觉泵，也让你理解好的直觉泵，更教你如何应用和制作直觉泵。





(1)　澳大利亚土著部落名。——编者注

(2)　最年轻的地质年代，开始于11 700年前。根据传统的地质学观点，全新世一直持续至今，但也有人提出，工业革命后应该另分为人类世。——编者注

(3)　使用多因素的变量协方差来进行分析，协方差常用于概率论和统计学中，是指两个变量的总体误差。——编者注

(4) 尼科洛·马基雅维利崇尚权术和谋略，并以主张“为达目的，不择手段”而著称。马基雅维利主义通常分为高马基雅维利主义和低马基雅维利主义；高马基雅维利主义者重视实效性，相信结果能替手段辩护，倾向于操纵他人以获得更多利益。——编者注

(5) 史蒂芬·平克是著名的心理学家，他的著作《当下的启蒙》已于2019年由湛庐文化引进、浙江人民出版社出版。——编者注

(6) 能在成员之间提供共同的陈述、目标和价值观的资源。——编者注

(7) 即博弈中的一种局面。在这种局面中，对于所有参与者而言，只要他人不改变策略，自身状况就无法得到改善。当每个参与者所能选择的策略种类都受到了限制，且混合策略被允许存在时，纳什均衡就会出现。——编者注

(8) 丹尼尔·丹尼特是著名的哲学家，他的著作《直觉泵和其他思考工具》已于2018年由湛庐文化引进、浙江教育出版社出版。——编者注

(9) 亲社会行为又叫利社会行为，是指符合社会希望，并对行为者本身无明显好处，而行为者却自觉自愿给行为受体带来利益的一类行为；行为出现偏差通常是由于不良的社会环境和教育方式，以及家庭功能失调所致。——编者注

(10) 理查德·道金斯的著作《基因之河》已于2019年由湛庐文化引进、浙江人民出版社出版。——编者注

(11) 约瑟夫·坎贝尔的著作《千面英雄》《英雄之旅》已分别于2016年、2017年由湛庐文化引进、浙江人民出版社出版。——编者注

(12) 原型是荣格所提出的集体无意识理论的关键词，准确地说是心理原型。它不是所经历之事留在大脑中的记忆表象，而是与生俱来的需要通过后天经验来体现的心理特质；若是无法通过意识来表现，就会化身为梦境、幻想和神经症症状。——编者注

(13) 通常用来表示消费者在消费中所获得的效用与所消费的商品组合之间的数量关系，以衡量消费满意度；其公理表达简单且规范，在形式上较好地体现了风险行为类型，但需要注意的是，预期效用理论中起关键作用的独立性公理假设并不总是成立。——编者注

(14) 艾莉森·高普尼克的著作《园丁与木匠》已于2019年由湛庐文化引进、浙江人民出版社出版。——编者注